OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: June 7, 2004, 15:33:46 ; Search time 3409 Seconds
(without alignments)
10908.866 Million cell updates/sec

Title: US-09-889-733B-1
Perfect score: 858
Sequence: 1 atggccggcctcggccaccc.........acaagaaggtagactcctga 858

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2171516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vi:*
29: em_vrt:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	858	100.0	858	6	BD251986	BD251986 Screen me
2	858	100.0	858	6	AX032835	AX032835 Sequence
3	851.6	99.3	1624	6	BC033680	BC033680 Homo sapi
4	851.6	99.3	1633	6	BD094073	BD094073 Shear str
5	851.6	99.3	1633	9	AB001915	AB001915 Homo sapi
6	851.6	99.3	4208	6	AX548046	AX548046 Sequence
7	739.4	86.2	3737	9	BC043235	BC043235 Homo sapi
8	717.2	83.6	1120	10	BC034505	BC034505 Mus muscu
9	715.6	83.4	3008	6	AX683109	AX683109 Sequence
10	715.6	83.4	3008	10	D86041	D86041 Rattus norv
11	579.6	67.6	1150	5	AB095027	AB095027 Gallus ga
12	555.4	64.7	3908	6	HSM808292	BX648145 Homo sapi
13	462.4	53.9	2289	10	AK128909	AK128909 Mus muscu
14	457.8	53.4	1344	5	BC056074	BC056074 Xenopus 1
15	381.8	44.5	3420	10	BC060266	BC060266 Mus muscu
16	328.8	38.3	2895	9	AK098405	AK098405 Homo sapi
17	300.2	35.0	104071	2	AL162260	AL162260 Homo sapi
18	300.2	35.0	158405	2	AC019261	AC019261 Homo sapi
19	300.2	35.0	168953	9	AL360219	AL360219 Human DNA
20	260.2	30.3	186434	2	AC126412	AC126412 Mus muscu
21	258.6	30.1	228940	2	AC126890	AC126890 Rattus no
22	216.2	25.2	1263	9	BC001435	BC001435 Homo sapi
23	216.2	25.2	1342	9	AK098671	AK098671 Homo sapi
24	216.2	25.2	1376	6	AX011712	AX011712 Sequence
25	216.2	25.2	1376	6	BD226323	BD226323 Pancreati
26	216.2	25.2	1709	9	AK026191	AK026191 Homo sapi
27	214.6	25.0	858	6	BD251987	BD251987 Screen me
28	214.6	25.0	858	6	AX032837	AX032837 Sequence
29	214.6	25.0	1351	9	AX779950	AX779950 Sequence
30	214.6	25.0	1351	9	AF087894	AF087894 Homo sapi
31	202.8	23.6	1310	10	BC003328	BC003328 Mus muscu
32	193.2	22.5	1270	10	AF004106	AF004106 Mus muscu
33	189.8	22.1	1228	6	AX780306	AX780306 Sequence
34	185.4	21.6	1397	6	AR379492	AR379492 Sequence
35	148	17.2	83946	9	HSA123M24	AL078459 Human DNA
36	122.4	14.3	58539	2	AC137903	AC137903 Mus muscu
37	122.4	14.3	200007	2	AC123684	AC123684 Mus muscu
38	119.8	14.0	212111	2	AC094482	AC094482 Rattus no
39	107.4	12.5	1377	3	AY060959	AY060959 Drosophil
40	97.8	11.4	186434	2	AC126412	AC126412 Mus muscu
41	93.2	10.9	212111	2	AC094482	AC094482 Rattus no
42	92.8	10.8	110000	2	AC144261_2	Continuation (3 of
43	92.8	10.8	304208	2	AC143028	AC143028 Macaca mu
44	85.6	10.0	626	11	BV064692	BV064692 S212P6014
45	85	9.9	420	6	AX898477	AX898477 Sequence

ALIGNMENTS

RESULT 1
BD251986
LOCUS   BD251986   858 bp   DNA   linear   PAT   17-JUL-2003
DEFINITION   Screen method.
ACCESSION   BD251986
VERSION   BD251986.1 GI:33061756
KEYWORDS   JP 2002535001-A/1.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 858)
  AUTHORS   Vallance,P.J.T., Leiper,J.M., Whitley,G.S.J. and Charles,I.G.
  TITLE   Screen method
  JOURNAL   Patent: JP 2002535001-A 1 22-OCT-2002;

Pred. No. is the number of results predicted by chance to have a

```
COMMENT    UNIVERSITY COLLEGE LONDON
OS   Homo sapiens (human)
PN   JP 2002535001-A/1
PD   22-OCT-2002
PF   26-JAN-2000 JP 2000596131
PR   26-JAN-1999 GB    9901705.5,04-JUN-1999 GB    9913066.8 PI
     PATRICK JOHN THOMPSON VALLANCE,JAMES MITCHELL LEIPER,GUY ST JI
     JOHN WHITLEY,
PI   IAN GEORGE CHARLES
PC   C12N15/09,A01K67/027,A61K31/198,A61K38/00,A61K45/00,A61P3/06,
     A61P9/02,
PC   A61P9/04,A61P9/10,A61P9/12,A61P13/12,A61P25/06,A61P25/18,  PC
     A61P25/28,
PC   A61P29/00,A61P31/04,A61P35/00,C07K16/40,C12N1/15,C12N1/19,  PC
     C12N1/21,
PC   C12N5/10,C12N9/78,C12Q1/02,C12Q1/34,G01N33/15,G01N33/50 PC
     ,C12N15/00,C12N5/00,
PC   A61K37/02
CC   Screen method
FH   Key          Location/Qualifiers
FT   CDS          (1)..(858).

ORIGIN
FEATURES
source    1..858
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"
ORIGIN

Query Match         100.0%;  Score 858;  DB 6;  Length 858;
Best Local Similarity  100.0%;  Pred. No. 8.1e-132;
Matches 858;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 ATGGCCGGCCTCGGCCACCCCTCGGCCGCCACCGCGCTGTGCGGGCG 60
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGGCCGGCCTCGGCCACCCCTCGGCCGCCACCGCGCTGTGCGGGCG 60

Qy   61 CTACCCGAGTCGCTCTGCCAGCACGCGCGCCAAGGGCGAGAAGCGCCAAGGGCGAGGAGTTGGACGTC 120
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTACCCGAGTCGCTCTGCCAGCACGCGCGCCAAGGGCGAGAAGCGCCAAGGGCGAGGAGTTGGACGTC 120

Qy  121 GCCCGCGCCGGAACGGCCAGCACCAGCTCTACGTGGCGTGCTGGGCAAGCTGGGGTG 180
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GCCCGCGCCGGAACGGCCAGCACCAGCTCTACGTGGCGTGCTGGGCAAGCTGGGGTG 180

Qy  181 CAGGTGGTGAGTGCGCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 240
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CAGGTGGTGAGTGCGCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 240

Qy  241 GCCGGTGGTGTGCGGAGGAACGGACGGCCCTCATCACCCGACCCGGGCCCAGCGGGAGAAG 300
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GCCGGTGGTGTGCGGAGGAACGGACGGCCCTCATCACCCGACCCGGGCCCAGCGGGAGAAG 300

Qy  301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360

Qy  361 GATGAAAATGCAACTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GATGAAAATGCAACTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420

Qy  421 GGCCTTTCCAAAAGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GGCCTTTCCAAAAGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480

Qy  481 TATGCAGTGTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TATGCAGTGTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540

Qy  541 GCTGGGCCTAACCTGATCGGTTGGGTCTAGTGAATCGCACGAGAAGGCCCCTTAAGATC 600
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GCTGGGCCTAACCTGATCCGTTGGGTCTAGTGAATCGCAACAGAAGGCCCCTTAAGATC 600

Qy  601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
```

```
Db  601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
          |||||||||||||||||||||||||||||||||||||||||||||||
Qy  661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCTGCGACCCCGGAA 720
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCTGCGACCCCGGAA 720
Qy  721 GAGTATCCAGAAAGTGCAAAGTTTATGAAGAACTGAAGGACCAATGCTGATCCCGTG 780
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGTATCCAGAAAGTGCAAAGTTTATGAAGAACTGAAGGACCAATGCTGATCCCGTG 780
Qy  781 AGCAATGTCTGAACTGGAAAAGTGGATGGGCTGCTCACCTGCTGCTCAGTTTAATTAAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AGCAATGTCTGAACTGGAAAAGTGGATGGGCTGCTCACCTGCTGCTCAGTTTAATTAAC 840
Qy  841 AAGAAGGTAGACTCCTGA 858
          ||||||||||||||||||
Db  841 AAGAAGGTAGACTCCTGA 858
```

```
RESULT 2
AX032835
LOCUS        AX032835               858 bp    DNA        linear   PAT 21-SEP-2000
DEFINITION   Sequence 1 from Patent WO0044888.
ACCESSION    AX032835
VERSION      AX032835.1  GI:10279809
KEYWORDS     .
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
  AUTHORS    Leiper,J.M., Whitley,G.S., Charles,I.G. and Vallance,P.J.
  TITLE      Screen method
  JOURNAL    Patent: WO 044888-A 1 03-AUG-2000;
             LEIPER JAMES MITCHELL (GB) ; WHITLEY GUY ST JOHN (GB) ; UNIV LONDON
             (GB) ; CHARLES IAN GEORGE (GB) ; VALLANCE PATRICK JOHN THOMPSON
             (GB)
FEATURES             Location/Qualifiers
     source          1..858
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1..858
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09947.1"
                     /db_xref="GI:10279810"
                     /db_xref="REMTREMBL:CAC09947"
                     /translation="MAGLGHPSAFGRATHAVVRALPESLCQHALRSAKEEVDVARAE
                     RQHQLYVGVLGSKLGLQVVELPADESLPDCVFVEDVAVVCEETALITRPGAPSRRKEV
                     DMKEALEKQLNIVEMKDRNATLDGGDVLFTGRRFFVGLSKRTNQRGAEILADTFKD
                     YAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALKIMQOMSDHRYDKLTVPDDI
                     AANCIYLNIPNKGHVLLHRFPEEYPESAKVYEKLKDHMLIPVSMSELEKVDGLILTCCS
                     VLINKKVDS"
ORIGIN

Query Match         100.0%;  Score 858;  DB 6;  Length 858;
Best Local Similarity  100.0%;  Pred. No. 8.1e-132;
Matches 858;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 ATGGCCGGCCTCGGCCACCCCTCGGCCGCCACCGCGCTGTGCGGGCG 60
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGGCCGGCCTCGGCCACCCCTCGGCCGCCACCGCGCTGTGCGGGCG 60

Qy   61 CTACCCGAGTCGCTCTGCCAGCACGCGCGCCAAGGGCGAGAAGCGCCAAGGGCGAGGAGTTGGACGTC 120
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTACCCGAGTCGCTCTGCCAGCACGCGCGCCAAGGGCGAGAAGCGCCAAGGGCGAGGAGTTGGACGTC 120

Qy  121 GCCCGCGCCGGAACGGCGCAGCACCAGCTCTACGTGGGGCTGGGGCTG 180
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GCCCGCGCCGGAACGGCGCAGCACCAGCTCTACGTGGGGCTGGGGCTG 180
```

```
Qy  181  CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCGGACTGCGTCTTCGTGGAGGACGTG  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCGGACTGCGTCTTCGTGGAGGACGTG  240

Qy  241  GCCGGTGGTGCGGAGGAGACGGGCCCTCATCACCCAGCACCCGGGCGGCCCGAGCCGGAG  300
         |||| |||||||||||||||||||||||||||||||  |||  ||||||||||||||||||
Db  241  GCCGTGGTGCGGAGGAGACGGGCCCTCATCACCCACCCGACCCGGGCGGCCCGAGCCGGAG 300

Qy  301  GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA  360

Qy  361  GATGAAAATGCAACTTAGTAGTGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTCGTG  420
         |||||||||  |||||||||||||||||||||||||  |||||||||||||||||| ||||
Db  361  GATGAAAATGCAACTTAGTGATGGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 420

Qy  421  GGCCTTCCAAAAGGACAAATCACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GGCCTTCCAAAAGGACAAATCACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC  480

Qy  481  TATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG  540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  TATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG  540

Qy  541  GCTGGGCCTAACCTGATCGGCAATTGGGTCTAGTGATCTGCACAGAAGGCCCTTAAGATC  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  GCTGGGCCTAACCTGATCGGCAATTGGGTCTAGTGATCTGCACAGAAGGCCCTTAAGATC  600

Qy  601  ATGCAACAGATGAGTGACCACCGCTACGACAACAAACTCACTGTGCCTGATGACATAGCAGCA  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  ATGCAACAGATGAGTGAGTGACCACCGCTACGACAACAAACTCACTGTGCCTGATGACATAGCAGCA 660

Qy  661  AACTGTATATATCTAAAATATCCCAACAAAGGCACGTCTTGCTGCACCGAACCCGGAA  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  AACTGTATATATCTAAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCGGAA  720

Qy  721  GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCCATATGTCGATCCCCGTG  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCCATATGTCGATCCCCGTG  780

Qy  781  AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAAATTAAC  840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAAATTAAC  840

Qy  841  AAGAAGGTAGACTCCTGA  858
         ||||||||||||||||||
Db  841  AAGAAGGTAGACTCCTGA  858
```

RESULT 3
BC033680

REFERENCE   1 (bases 1 to 1624)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I. Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2 (bases 1 to 1624)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,J.W., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 68 Row: m Column: 1
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 6912327.

FEATURES   Location/Qualifiers
     source   1..1624
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:45161 IMAGE:5189970"
                     /tissue_type="Colon, Kidney, Stomach, adult, whole pooled"
                     /clone_lib="NIH_MGC_116"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene   1..1624
                     /gene="DDAH1"
                     /note="synonym: DDAH"
                     /db_xref="LocusID:23576"
                     /db_xref="MIM:604743"
     CDS   100..957
                     /codon_start=1
                     /product="DDAH1 protein"
                     /protein_id="AAH33680.1"
                     /db_xref="GI:21707415"
                     /db_xref="LocusID:23576"
                     /translation="MAGLGHPAAFGRATHAVVRALPESLGQHALRSAKGEVDVARAE
                     RQHQLYVGVLGSKLQVVELPADESLPDCVFVRDVAVVCETALITRPGAPSRRKEV
                     DMMKBALEKLQLNIVEMKDENATLDGDVLFTGREFPVGLSKRTNORGAEILADTFKD
                     YAVSTVPADGLHLKSFCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLITVPDDI
                     AANCIYLNIPNKGHVLLHRTPEEYPESAKYYEKLKDHMLIPVSMSLEKVDGLLTCCS
                     VLINKKVDS"
     misc_feature   133..936
                     /note="Amidinotransf; Region: Amidinotransferase. This

RESULT 4

LOCUS   BD094073   1633 bp   DNA   linear   PAT 27-AUG-2002
DEFINITION Shear stress-responsive DNAs.
ACCESSION   BD094073
VERSION   BD094073.1 GI:22639661
KEYWORDS   WO 0125427-A/34.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1633)
  AUTHORS   Nojima,H., Yoshisue,H., Obayashi,M., Ota,T., Kawabata,A.,
            Sakurada,K., Kuga.T., Sekine,S., Nakamura,Y. and Sugan,S.
  TITLE   Shear stress-responsive DNAs
  JOURNAL   Patent: WO 0125427-A 34 12-APR-2001;
            KYOWA HAKKO KOGYO CO LTD,HIROSHI NOJIMA,HAJIME YOSHISUE, MASAYA
            OBAYASHI, TOSHIO OTA,AYAKO KAWABATA,KAZUHIRO SAKURADA,TETSURO KUGA,
            SUSUMU SEKINE, YUSUKE NAKAMURA,SUMIO SUGANO
COMMENT   OS   Homo sapiens (human)
            PN   WO 0125427-A/34
            PD   12-APR-2001
            PF   02-OCT-2000 WO 2000JP006840
            PR   01-OCT-1999 JP 99P 280976
            PI   HIROSHI NOJIMA,HAJIME YOSHISUE,MASAYA OBAYASHI, TOSHIO OTA, PI
            AYAKO KAWABATA,
            PI   KAZUHIRO SAKURADA,TETSURO KUGA, SUSUMU SEKINE,YUSUKE NAKAMURA,
            PI   SUMIO SUGANO
            PC   C12N15/12,C07K14/435,C07K16/18,C12P21/02,C12Q1/68,A61K38/00,
            PC   A61K39/395,
            PC   A61K48/00,A61P9/10,G01N33/50,G01N33/53
            CC
            FH   Key   Location/Qualifiers
            FT   CDS   (323)..(1177).
FEATURES   Location/Qualifiers
     source   1..1633
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   99.3%; Score 851.6; DB 6; Length 1633;
Best Local Similarity 99.5%; Pred. No. 9.1e-131;
Matches 854; Conservative 0; Mismatches 0; Indels 4; Gaps 0;

Qy   1 ATGGGCCGGCCTCGGCCACCCCTCCGGCCTTCGGCCGCGGCCACCCACGCCGTGGTGCGGGCG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323 ATGGGCCGGGCTCGGCCACCCCTCCGGCCTTCGGCCGCGGCCACCCACGCCGTGGTGCGGGCG 382

Qy   61 CTACCCAGTCGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGAAGGTGGACGTC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   383 CTACCCAGTCGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGAAGGTGGACGTC 442

Qy   121 GCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGGCGTGCTGGGCAGCTGGGCTG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   443 GCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGGCGTGCTGGGCAGCTGGGGCTG 502

Qy   181 CAGGTGGTGGAGCTGCCGCGACGAGCCTTCCGGACTCGTCTTCGTGTGGAGACGTG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   503 CAGGTGGTGGAGCTGCCGCGGCCGACGAGAGCCTTCCGGACTCGTCTTCGTGTGGAGACGTG 562

Qy   241 GCCGTGGTGTGCCGAGGAGACGGCCTCATCACCCGACCCGGGCGCGCCGAGCCGGAGGAAG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   563 GCCGTGGTGTGCCGAGGAGACGGCCTCATCACCCGACCCGGGCGCGCCGAGCCGGAGGAAG 622

Qy   301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   623 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 682

Qy   361 GATGAAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

---

family contains glycine (EC:2.1.4.1) and inosamine
(EC:2.1.4.2) amidinotransferases, enzymes involved in
creatine and streptomycin biosynthesis respectively. This
family also includes arginine deiminases, EC:3.5.3.6.
These enzymes catalyse the reaction: arginine + H2O <=>
citrulline + NH3. Also found in this family is the
Streptococcus anti tumor glycoprotein"
/db_xref="CDD:pfam02274"

ORIGIN

Query Match   99.3%; Score 851.6; DB 9; Length 1624;
Best Local Similarity 99.5%; Pred. No. 9.1e-131;
Matches 854; Conservative 0; Mismatches 0; Indels 4; Gaps 0;

Qy   1 ATGGGCCGGCCTCGGCCACCCCTCCGGCCTTCGGCCGCGGCCACCCACGCCGTGGTGCGGGCG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   100 ATGGGCCGGGCTCGGCCACCCCTCCGGCCTTCGGCCGCGGCCACCCACGCCGTGGTGCGGGCG 159

Qy   61 CTACCCAGTCGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGAAGGTGGACGTC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   160 CTACCCAGTCGCTCGGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGAAGGTGGACGTC 219

Qy   121 GCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGGCGTGCTGGGCAGCTGGGGCTG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   220 GCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGGCGTGCTGGGCAGCTGGGGCTG 279

Qy   181 CAGGTGGTGGAGCTGCCGCGGACGAGCCTTCCGGACTCGTCTTCGTGTGGAGACGTG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   280 CAGGTGGTGGAGCTGCCGCGGACGAGCCTTCCGGACTCGTCTTCGTGTGGAGACGTG 339

Qy   241 GCCGTGGTGTGCCGAGGAGACGGCCTCATCACCCGACCCGGGCGCGCCGAGCCGGAGGAAG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   340 GCCGTGGTGTGCCGAGGAGACGGCCTCATCACCCGACCCGGGCGCGCCGAGCCGGAGGAAG 399

Qy   301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   400 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 459

Qy   361 GATGAAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   460 GATGAAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 519

Qy   421 GGCCTTTCCAAAAGGACAAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   520 GGCCTTTCCAAAAGGACAAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 579

Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   580 TATGCAGTCTCCACAGTGCCAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATG 639

Qy   541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAAATCTGCACAGAGAAGGCCCTTAGATC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   640 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAAATCTGCACAGAGAAGGCCCTTAGATC 699

Qy   601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   700 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 759

Qy   661 AACTGTATATATCTAAATATCCCAACAAAAGGCACGTCTTGCTGCACCCGAACCCGGAA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   760 AACTGTATATATCTAAATATCCCAACAAAAGGGCACGTCTTGCTGCACCCGAACCCGGAA 819

Qy   721 GAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   820 GAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG 879

Qy   781 AGCATGTCTGAACTGGAAAAGTTGGATGGGCTTGCTCACCTGCTGCAGTTTAATTAAC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   880 AGCATGTCTGAACTGGAAAAGTTGGATGGGCTTGCTCACCTGCTGCAGTTTAATTAAC 939

Qy   841 AAGAAGGTAGACTCCTGA 858
         ||||||||||||||||||
Db   940 AAGAAGGTAGACTCCTGA 957

```
/protein_id="BAA37117.1"
/db_xref="GI:4160666"
/translation="MAGLGHPAFGRATHAVVRALPESLGQHALRSAKGEVVDVARAE
RQHQLYVGVLGSKLGLQVVELPADESLPDCVFEDVAVVCEETALITRPGAPSRRKEV
DMMKEALEKLQLNIVEMKDENATLDGGDVLPTGREFVVGLSKRTNQRGAEILADTFKD
YAVSTVPVADGLHLKS?CSMAGPNLIAIGSSSASQKALKIMCQMSDHRYDKLIVPDDI
AANCIYLNIPNKGHVLLHRTPEYPESAKVYEKLKDHMLIPVSMSELEKVDGLLTCCS
VLINKKVDS"

ORIGIN

Query Match         99.3%;  Score 851.6;  DB 9;  Length 1633;
Best Local Similarity  99.5%;  Pred. No. 9.1e-131;
Matches 854;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy      1 ATGGCCGGCTCGGCCACCCTCCGGCGCCACCCCGCCGGGTCGCGGGCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 ATGGCCGGCTCGGCCACCCTCCGGCGCCACCCCGCCGGGTCGCGGGCG 382

Qy     61 CTACCCGAGTCGCTCTGCCAGCACCGCCGGCTCGAGAGCGCCAAGGGCGCAGGAGGTGGACGTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 CTACCCGAGTCGCTCTGCCAGCACCGCCGGCTCGAGAGCGCCAAGGGCGCAGGAGGTGGACGTC 442

Qy    121 GCCCGCGGAACGGCAGCAGCTCACGTGGGCGTGCTGGGCAAGCTCGGGCGTG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 GCCCGCGGAACGGCAGCAGCTCACGTGGGCGTGCTGGGCAAGCTCGGGCGTG 502

Qy    181 CAGGTGGTGGAGCTGCCGGCGACGAGAGCCTTCCGGACGTGCGTCTTCGTGGAGGACGTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 CAGGTGGTGGAGCTGCCGGCGACGAGAGCCTTCCGGACGTGCGTCTTCGTGGAGGACGTG 562

Qy    241 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGGCCCGAGCCCGGAGGAAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGGCCCGAGCCCGGAGGAAG 622

Qy    301 GAGGTTGACATGATGAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 GAGGTTGACATGATGAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAA 682

Qy    361 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 742

Qy    421 GGCCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 GGCCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 802

Qy    481 TATGCAGTCTCCACAGTGCCAGTTGGTCAGTGCAGATGGGTTGCATTTGAAGAGTTTCTGCGCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    803 TATGCAGTCTCCACAGTGCCAGTTGGTCAGTGCAGATGGGTTGCATTTGAAGAGTTTCTGCGCATG 862

Qy    541 GCTGGGCCTAACCTGATCGCCAATTGGGCTCAGTGAATCTGCACAGAAGGCCCTAAGATC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    863 GCTGGGCCTAACCTGATCGCCAATTGGGCTCAGTGAATCTGCACAGAAGGCCCTAAGATC 922

Qy    601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    923 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 982

Qy    661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGACCCCGGAA 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    983 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGACCCCGGAA 1042

Qy    721 GAGTATCCAGAAAGTGCAAAGGTTTATGAAGAACTGAAGGACCATATGTCGATCCCGTG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1043 GAGTATCCAGAAAGTGCAAAGGTTTATGAAGAACTGAAGGACCATATGTCGATCCCGTG 1102

Qy    781 AGCATGTCTGAAAAGTGGATGGCTGCTCACCTGCTCAGTTTTAATTAAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1103 AGCATGTCTGAAAAGTGGATGGCTGCTCACCTGCTCAGTTTTAATTAAC 1162

Qy    841 AAGAAGGTAGACTCCTGA 858
          ||||||||||||||||||
Db   1163 AAGAAGGTAGACTCCTGA 1180
```

```
Db    683 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 742

Qy    421 GGCCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 480

Db    743 GGCCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 802

Qy    481 TATGCAGTCTCCACAGTGCCAGTTGGTCAGTGCAGATGGGTTGCATTTGAAGAGTTTCTGCGCATG 540

Db    803 TATGCAGTCTCCACAGTGCCAGTTGGTCAGTGCAGATGGGTTGCATTTGAAGAGTTTCTGCGCATG 862

Qy    541 GCTGGGCCTAACCTGATCGCCAATTGGGCTCAGTGAATCTGCACAGAAGGCCCTAAGATC 600

Db    863 GCTGGGCCTAACCTGATCGCCAATTGGGCTCAGTGAATCTGCACAGAAGGCCCTAAGATC 922

Qy    601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660

Db    923 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 982

Qy    661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGACCCCGGAA 720

Db    983 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGACCCCGGAA 1042

Qy    721 GAGTATCCAGAAAGTGCAAAGGTTTATGAAGAACTGAAGGACCATATGTCGATCCCGTG 780

Db   1043 GAGTATCCAGAAAGTGCAAAGGTTTATGAAGAACTGAAGGACCATATGTCGATCCCGTG 1102

Qy    781 AGCATGTCTGAAAAGTGGATGGCTGCTCACCTGCTCAGTTTTAATTAAC 840

Db   1103 AGCATGTCTGAAAAGTGGATGGCTGCTCACCTGCTCAGTTTTAATTAAC 1162

Qy    841 AAGAAGGTAGACTCCTGA 858

Db   1163 AAGAAGGTAGACTCCTGA 1180
```

```
RESULT 5
AB001915
LOCUS       AB001915              1633 bp    mRNA      linear    PRI 21-SEP-2000
DEFINITION  Homo sapiens mRNA for NG,NG-dimethylarginine
            dimethylaminohydrolase, complete cds.
ACCESSION   AB001915
VERSION     AB001915.1  GI:4160665
KEYWORDS    NG,NG-dimethylarginine dimethylaminohydrolase.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
AUTHORS     Kimoto,M., Miyatake,S., Sasagawa,T., Yamashita,H., Okita,M.,
            Oka,T., Ogawa,T. and Tsuji,H.
TITLE       Purification, cDNA cloning and expression of human
            NG,NG-dimethylarginine dimethylaminohydrolase
JOURNAL     Eur. J. Biochem. 258 (2), 863-868 (1998)
MEDLINE     99089653
PUBMED      9874257
REFERENCE   2  (bases 1 to 1633)
AUTHORS     Kimoto,M.
TITLE       Direct Submission
JOURNAL     Submitted (15-MAR-1997) Masumi Kimoto, Okayama Prefectural
            University, Faculty of Health and Welfare Science, Department of
            Nutritional Science; Kuboki 111, Soja, Okayama 719-1197, Japan
            (E-mail:kimoto@fhw.oka.-pu.ac.jp, Tel:0866-94-2158,
            Fax:0866-94-2158)
FEATURES             Location/Qualifiers
     source          1..1633
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="Kidney"
     CDS             323..1180
                     /EC_number="3.5.3.18"
                     /codon_start=1
                     /product="NG,NG-dimethylarginine dimethylaminohydrolase"
```

RESULT 6
AX548046
LOCUS   AX548046   4208 bp   DNA   linear   PAT 26-NOV-2002
DEFINITION Sequence 22 from Patent WO02066654.
ACCESSION   AX548046
VERSION   AX548046.1 GI:25813142
KEYWORDS
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Astromoff,A., Au-Young,J., Baughn,M.R., Ding,L., Duggan,B.M.,
            Forsythe,I.J., Gietzen,K.J., Griffin,J.A., Lee,E.A., Lu,Y.,
            Richardson,T.W., Ring,H.Z., Sanjanwala,M.M., Swarnakar,A.,
            Walia,N.K., Warren,B.A., Xu,Y., Yue,H. and Zebarjadian,Y.
  TITLE   Drug metabolizing enzymes
  JOURNAL   Patent: WO 02066654-A 22 29-AUG-2002;
            Incyte Genomics, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..4208
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7671089CB1"
ORIGIN

Query Match   99.3%; Score 851.6; DB 6; Length 4208;
Best Local Similarity   99.5%; Pred. No. 8.9e-131;
Matches 854; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   1 ATGGCCGGCTCGGCGCCACCCCTCGGCCTTCGGCCGCGGCCACCCACGCCGTGGTGCGGCG   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ATGGGCCGGGCTCGGCGCCACCCCGCCGGCCACCCACGCCGTGGTGCGGGCG   427

Qy   61 CTACCCGAGTGCTCTGCCGCACGGCGCCAAGGGCGAGGAGGTGGACGTC   120
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 CTACCCGAGTGCTCTGGCCAGCACGGCGCCAAGGGCGAGGAGGTGGACGTC   487

Qy   121 CCCGCGCGGAACGGCAGCAGCACCAGCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 GCCCGCGCGGAACGGCAGCAGCACCAGCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG   547

Qy   181 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGGAGGACGTG   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGGAGGACGTG   607

Qy   241 GCCGTGGTGTCGGAGAGACGGCCCTCATCACCCGACACCCCGGGGCCCGAGCCGGAGGAG   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 GCCGTGGTGTCGGAGAGACGGCCCTCATCACCCGACACCCCGGGGCCCGAGCCGGAGGAG   667

Qy   301 GAGGTGACATGATGAAAGAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 GAGGTTGACATGATGAAAGAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA   727

Qy   361 GATGAAAATGCAACTTAGATGGCCGAGATGTTTATTCACACAGGCCAGAGAATTTTTGTG   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 GATGAAAATGCAACTTAGATGGCCGAGATGTTTATTCACACAGGCCAGAGAATTTTTGTG   787

Qy   421 GGCCTTTCCAAAAGGACAAATCAACCAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 GGCCTTTCCAAAAGGACAAATCAACCAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC   847

Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTCTGCAGCATG   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTCTGCAGCATG   907

Qy   541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTAAGATC   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC   967

Qy   601 ATGCAACAGATGAGTGACCACCGCTAGACAACTCACTGTGCTGATGACATAGCAGCA   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 ATCCAACAGATGAGTGACCACCGCTAGACAACTCACTGTGCTGATGACATAGCAGCA   1027

Qy   661 AACTGTATATCTAAATATCCCAACAAAGGCACGTCTTGCTGCACCGAACCCCGAA   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 AACTGTATATCTAAATATCCCAACAAAGGCACGTCTTGCTGCACCGAACCCCGAA   1087

Qy   721 GAGTATCCAGAAGTGCAAAGGTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 GAGTATCCAGAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG   1147

Qy   781 AGCATGTCTGAACTGGAAAAGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 AGCATGTCTGAACTGGAAAAGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC   1207

Qy   841 AAGAAGGTAGACTCCTGA   858
          ||||||||||||||||||
Db   1208 AAGAAAGTAGACTCCTGA   1225

RESULT 7
BC043235
LOCUS   BC043235   3737 bp   mRNA   linear   PRI 30-SEP-2003
DEFINITION Homo sapiens dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
            clone IMAGE:5295723), partial cds.
ACCESSION   BC043235
VERSION   BC043235.2 GI:34783628
KEYWORDS
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 3737)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2 (bases 1 to 3737)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (09-JAN-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
  COMMENT   On Sep 16, 2003 this sequence version replaced gi:28175756.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D.,Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,

```
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 76 Row: j Column: 15.

FEATURES             Location/Qualifiers
     source          1..3737
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5295723"
                     /tissue_type="Testis"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note=Vector: pBluescript"
     gene            <1..3737
                     /gene="DDAH1"
                     /note="synonym: DDAH"
                     /db_xref="LocusID:23576"
                     /db_xref="MIM:604743"
     CDS             <1..741
                     /gene="DDAH1"
                     /codon_start=1
                     /product="DDAH1 protein"
                     /protein_id="AAH43235.2"
                     /db_xref="GI:34783629"
                     /db_xref="LocusID:23576"
                     /db_xref="MIM:604743"
                     /translation="VARAERQHQLYVGVLGSKLGLQVVELPADESLPDCVFVEDVAVV
                     CEETALITRPGAPSRRKEVDMKEALEKLQLNIVEMKDENATLDGGDVLFTGREFFVG
                     LSKRTNQRGAEILADTFKDYAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALK
                     IMQQMSDHRYDKLTVPDDIAANCIYLNIPNKGHVLLHRTPEYPESAKVYEKLKDHML
                     IPVSMSELEKVDGLLTCCSVLINKKVDS"
     misc_feature    1..720
                     /gene="DDAH1"
                     /note="Amidinotransf; Region: Amidinotransferase. This
                     family contains glycine (EC:2.1.4.1) and inosamine
                     (EC:2.1.4.2) amidinotransferases, enzymes involved in
                     creatine and streptomycin biosynthesis respectively. This
                     family also includes arginine deiminases, EC:3.5.3.6.
                     These enzymes catalyse the reaction: arginine + H2O <=>
                     citrulline + NH3. Also found in this family is the
                     Streptococcus anti tumor glycoprotein"
                     /db_xref="CDD:pfam02274"
ORIGIN

Query Match       86.2%;  Score 739.4;  DB 9;  Length 3737;
Best Local Similarity 99.9%;  Pred. No. 2.8e-112;
Matches  740;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   118 GTCGCCCGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGG 177
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GTCGCCCGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGG 60

Qy   178 CTGCAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACGCGTCTTCGTGGAGGAC 237
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTGCAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACGCGTCTTCGTGGAGGAC 120

Qy   238 GTGGCCGTGGTGTGCGAGGAGACGGCGCCCTCATCACCCGACCCGGGGCGGCCGAGCCGAGG 297
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGGCCGTGGTGTGCGAGGAGACGGCGCCCTCATCACCCGACCCGGGGCGGCCGAGCCGAGG 180

Qy   298 AAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATG 357
             |||||||||||||||||||| |||||||||||||||||||||||||||||||| |||||
Db   181 AAGGAGGTTGACATGATGAAGAAGAAGCATTAGAAAAACTTCAGCTCAATATATAGAGATG 240

Qy   358 AAAGATGAAAATGCAACTTTAGATGCGGAGATGTTTTATTCACAGCAGAATTTTT 417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AAAGATGAAAATGCAACTTTAGATGCGGAGATGTTTTATTCACAGCAGAATTTTT 300

Qy   418 GTGGGCCTTCAAAAGGACAAATCACGAGGTGCTGAAATCTTGGCTGATACTTTTAAG 477
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GTGGGCCTTCAAAAGGACAAATCACGAGGTGCTGAAATCTTGGCTGATACTTTTAAG 360

Qy   478 GACTATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTGGAAGAGTTTCTGCAGC 537
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GACTATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTGGAAGAGTTTCTGCAGC 420

Qy   538 ATGGCTGGGCCTTAACCTGATCCCAATTGGGTCTAGTGAATCTGCACAGAAGAGCCCTTAAG 597
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ATGGCTGGGCCTTAACCTGATCCCAATTGGGTCTAGTGAATCTGCACAGAAGAGCCCTTAAG 480

Qy   598 ATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCA 657
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCA 540

Qy   658 GCAAACTGTATATATCCCAACAAAGGGCACGTCTTGTGCACCGAACCCG 717
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCAAACTGTATATATCCCAACAAAGGGCACGTCTTGTGCACCGAACCCG 600

Qy   718 GAAGAGTATCCAGAAAGTGCAAAGGTTTATGAGAAATCGAAGGACCATATCTGATCCCC 777
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GAAGAGTATCCAGAAAGTGCAAAGGTTTATGAGAAATCGAAGGACCATATCTGATCCCC 660

Qy   778 GTGGAGCATGTCTGAACTGGAAAAGTGGATGGGCTGCTCACCTGCTCAGTTTTAATT 837
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GTGGAGCATGTCTGAACTGGAAAAGTGGATGGGCTGCTCACCTGCTCAGTTTTAATT 720

Qy   838 AACAAGAAGGTAGACTCCTGA 858
             |||||||| ||||||||||||
Db   721 AACAAGAAAGTAGACTCCTGA 741
```

```
Qy     1 ATGGCCGGCCTCGGCCACCCCTTCGGCCTTCGGCCTCGGCGGCCACCCACGCGTGTGCGGGCG   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 ATGGCCGGCCTCGGCCACCCCTTCGGCCTTCGGCCTCGGCGGCCACCCACCGCGTGTGCGGGGCT  181

Qy    61 CTACCCAGTCGCTCTGCCAGCACGCGGCTGCTGAGAAGCGGCCAAGGGCGAGGAGGTGGACGTC  120
         ||||||| |||| ||||||||||||||||  |||||||||||||||||||||||||||||||
Db   182 CCGCCCAGTCCCTGTGCCGCCACGCGCTGCGAGGGCGCTGCGCAGGGCGAGGAGGTGGACCTG  241

Qy   121 GCCCGGCCGGACGGCAGCAGCCAGCTCTACGTGGGCGGTCTGTGGGCAGCAAGCTGGGGGCTG  180
         ||||||||| |||  |||||||| |||||||||||||||||||||||||||||||||||||
Db   242 GCTCGGCCGGAGCGGCCAGCACGAGCTCTACGTGGGCGGTCTGGGCCAGCAAGCTGGGGGCTG  301

Qy   181 CAGGTGGTGGAGCTGCCGGCGGCCACGGAGAGCCTTCCGGACTGGCTCTTCGTGGAGGACGTG  240
         ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db   302 CAGGTGGTGGAGCTGCCCGGCGCCCACGAGAGCCGCTGCGTGTCGGTGGAGGAGGACGTG  361

Qy   241 GCCGTGGTGTGCGGAGGAGACGGCCCCTCATCACCCCGGGGCGCCGAGCCGGAGGAAG  300
         ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   362 GCCGTCGTGTGCGGAGGAGACGGCCCCTCATCACCCCGGGGGCGCCCAGCCGCCAGGAAG  421

Qy   301 GAGGTTGACATGATGAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA  360
         |||||||||||||||||||||||| ||||||||||||||| ||||||||||||||||||
Db   422 GAGGTTGACATGATGATGAAGAGAGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAA  481

Qy   361 GATGAAAATGCAACTTTAGATGGCGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG  420
         |||||||||||||||||||||||| ||| || ||||||||||||||||||||||||||
Db   482 GATGAAAATGCAACTTTCGATGGTGGGGACGTCCATTCACAGGCAGAGAATTTTTTGTG  541

Qy   421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC  480
         ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   542 GGCCTTTCCAAAAGGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC  601

Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAAAGAGTTTCTGCAGCATG  540
         |||||||||||| |||||||||||||||| ||||||||||||||||||| ||||||||||
Db   602 TACGCAGTTCTACAGTCCCTGTGGCCATTCTTGCATTTAAAGAGTTTCTGCAGCATG  661

Qy   541 GCTGGGCCTAACCTGATCGCAATTGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC  600
         ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   662 GCCGGACCCAACCTGATTGCAATAGGTCCAGCGAATCTGCACAGAGGCCCTCAAGATC  721

Qy   601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA  660
         ||||| |||| ||| ||||||||||||| |||||||||||||||||||||||||||||||
Db   722 ATGCAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCCGACCATGGCCCGCC  781

Qy   661 AACTGTATATATCTAAATATCCCCAACAAAGGGCACGTCTTGCTGCACGGAACCCCCGAA  720
         |||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||||
Db   782 AACTGTATATATCTAAATATCCCCGACAAAGGGCATGTCTTCGCTGCACCGAACCCCAGAA  841

Qy   721 GAGTATCCAGAAGTGCAAAGGTTTATGAGAAACTGAAGACCACCATATGCTGATCCCGTG  780
         |||| |||||||| |||||||||||| |||||||||||||||||||||||||||||||||
Db   842 GAGTACCCAGAAGCGCAAAGGTTCATGAGAAACTCAAGGACCATCTACTGATCCCGTG  901

Qy   781 AGCATGTCTGAACTGGAAAAGGTGGATCGGCTGCTCACCTGCTTGCTCAGTTTTAATTAAC  840
         ||||||||||| ||||||| ||||||||||| |||||||||||||||||||||| ||||||
Db   902 AGCAACTCGGAGATGGAAAAGGTGGACGGCTTGCTCACCTGCTTGCTCTGTTTTATTAAC  961

Qy   841 AAGAAGGTAGACTCCTGA  858
         |||| ||||||||||||||
Db   962 AAGAAGATAGACTCCTGA  979
```

RESULT 9

AX683109

LOCUS AX683109 3008 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 83 from Patent EP1279744.
ACCESSION AX683109
VERSION AX683109.1 GI:29370146
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 1120)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2002) National Institutes of Health, Mammalian
    Gene Collection (MGC), Cancer Genomics Office, National Cancer
    Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
    USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
    Email: cgapbs-r@mail.nih.gov
    Tissue Procurement: Jeffrey E. Green, M.D.
    cDNA Library Preparation: Life Technologies, Inc.
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
    DNA Sequencing by: Baylor College of Medicine Human Genome
    Sequencing Center
    Center code: BCM-HGSC
    Web site: http://www.hgsc.bcm.tmc.edu/cdna/
    Contact: amg@bcm.tmc.edu
    Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
    Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
    A.N., Gibbs, R.A.

    Clone distribution: MGC clone distribution information can be found
    through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
    Series: IRAK Plate: 41 Row: o Column: 4
    This clone was selected for full length sequencing because it
    passed the following selection criteria: Hexamer frequency ORF
    analysis; Similarity but not identity to protein.
FEATURES Location/Qualifiers
  source 1..1120
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="FVB/N"
    /db_xref="taxon:10090"
    /clone="MGC:30267 IMAGE:5134831"
    /tissue_type="Liver, normal, 5 month old male mouse."
    /clone_lib="NCI_CGAP_Li9"
    /lab_host="DH10B"
    /note="Vector: pCMV-SPORT6"
  gene 1..1120
    /gene="Ddah1"
    /note="synonym: 2510015N06Rik"
    /db_xref="LocusID:69219"
    /db_xref="MGI:1916469"
  CDS 122..979
    /codon_start=1
    /product="dimethylarginine dimethylaminohydrolase 1"
    /protein_id="AAH34505.1"
    /db_xref="GI:21961624"
    /db_xref="LocusID:69219"
    /translation="MAGLGHPSAFGRATHAVVRAPPESLCRHALRRSQGEVDFARAE
    RQHELYVGVLGSKIGLQVVQLPADESLPDCVFBDVAVVCEETALITRPGAPSRRKEV
    DMMKEALBKLQLNIVEMCDENATLDGGDVLFTGREFFVGLSXRTNORGABILADTFKD
    YAVSTVPVADSLHLKSFCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDM
    AANCIYLNIPSKGHVLLHRTPEEYPESAKVYEKLKDHLLLIPVSNSEMEKVDGLLTCCS
    VFINKKIDS"
  misc_feature 155..958
    /note="Amidinotransf; Region: Amidinotransferase. This
    family contains glycine (EC:2.1.4.1) and inosamine
    (EC:2.1.4.2) amidinotransferases, enzymes involved in
    creatine and streptomycin biosynthesis respectively. This
    family also includes arginine deiminases, EC:3.5.3.6.
    These enzymes catalyse the reaction: arginine + H2O <=>
    citrulline + NH3. Also found in this family is the
    Streptococcus anti tumour glycoprotein"
    /db_xref="CDD:pfam02274"
ORIGIN

Query Match 83.6%; Score 717.2; DB 10; Length 1120;
Best Local Similarity 89.7%; Pred. No. 1.3e-108;
Matches 770; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

REFERENCE   1
AUTHORS   Brooksbank,R.A., Dixon,A.K., Lee,K. and Pinnock,R.D.
TITLE   Identification and use of molecules implicated in pain
JOURNAL   Patent: EP 1279744-A 83 29-JAN-2003;
             WARNER-LAMBERT COMPANY (US)
FEATURES   Location/Qualifiers
     source   1..3008
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
                     /note="N-G, N-G dimethylarginine dimethylaminohydrolase"
ORIGIN

Query Match   83.4%;   Score 715.6; DB 6; Length 3008;
Best Local Similarity   89.6%;   Pred. No. 2.4e-108;
Matches   769; Conservative   0; Mismatches   89; Indels   0; Gaps   0;

Qy   1 ATGGCCGGCCTCGGCCACCCCTCGGCCCTTCGGCCGGCCACCCAGCCGTTGGTGCGGGCG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   432 ATGGCCGGCCTCAGCCACCCCTCGGTCTTCGGCCGGCCACCCAGCCGTTGGTGCGGGGCT 491

Qy   61 CTACCCGAGTCGCTCGCCAGCACGGCAGCACCAGCTCTACGTGGGCGGTGGGCAGCAGGTC 120
          ||||||||||||||||||||||| || |||| ||||| |||||| ||||||||||||||||
Db   492 CCGCCCGAGTCCCTGTCGCCGCACGGCGCTCGAGGGCGTCCCAGGGCGAAGAGGTGGATTTC 551

Qy   121 GCCGCGCGGACACGGCAGCAGCACCAGCTCTACGTGGGCGGTGTGGGCAGCAAGCTGGGCTG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   552 GCTCGCGCTGAGCGCCGCAGCACCAGCTCTACGTGGGCCGTGTGGGCAGCAAGCTGGGCGTG 611

Qy   181 CAGGTGGTGGAGACTGCCGGCCGACGAGAGCGCCTTCCGGACTGCGTCTTCGTTGGAGGACGTG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   612 CAGGTGGTGCAGCTGCCGGCCGACGAGAGCCTGCCTTGACTGCGTGTTCGTTGGAGGACGTG 671

Qy   241 GCCGTGGTGTGCGGAGGAGACAGGGCCCTCATCACCCGACCCGGGGGCCGAGCCGAGGAAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   672 GCCGTGGTGTGCGGAGGAGACCAGGGCCCCTAGCCGACGGCCGCCAGGAAG 731

Qy   301 GAGGTTGACAATGATGAAAGAAGCATTAGAAAAACTTCAGTCAATATAGTAGTAGAGATGAAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   732 GAGGTTGACAATGATGAAAGAAGCCTTTGAAAAACTTCAGTCCAACATAGTAGTAGAGATGAAA 791

Qy   361 GATGAAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 420
          |||||||||||||||||||| || ||||||| |||| |||||||||||||||||||||
Db   792 GATGAAAATGCAACCTTAGATGTGTGGGACGTCCTATTCACAGGCAGAGAGTTTTTTGTG 851

Qy   421 GGCCTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   852 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGAC 911

Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCGCAGATGGTTGCATTGAAAGAGTTTCTGCGCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   912 TACGCAGTTCCACAGTCCCGTGCCCGGCGATTCTTGCATTTAAAGAGTTTCTGCGCATG 971

Qy   541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   972 GCTGGCCCCAACCTGATCGCAATAGGGTCCAGTGAATCTGCACAGAAGGCCCTCAAGATC 1031

Qy   601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTCCTGATGACATAGCAGCA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1032 ATGCAACAGATGAGTGACCACCGTTACGACAAGCTCACTGTCCTGATGACAAGGCCGCC 1091

Qy   661 AACTGTATATATCTAAATATCCCAACAAAGGCCAGTCTTGCTGACACCGGAA 720
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   1092 AACTGTATATATTTAAATATCCCAGCAAAGGGCATGTTCTTGTCTGCTGCCGAACCCCAGAA 1151

Qy   721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCGTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1152 GAGTACCCAGAAAGCGCAAAGGTTTATGAGAAGCTCAAGGACCATCTCACTGATCCCTGTG 1211

Qy   781 AGCATGTCTGAACTGGAAAAGTGGATGGGCTGCTCACCTGCTCCAGTTTAATTAAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1212 AGCAATTCTGAGATGGAAAAAGTGGATGGACCGGCTTGCTCACCTGCTCCGTTTTTATTAAC 1271

Qy   841 AAGAGGTAGACTCCTGA   858
          |||||||||||||||||
Db   1272 AAGAGAACAGACTCTTGA 1289

RESULT 10
D86041
LOCUS   D86041   3008 bp   mRNA   linear   ROD 06-FEB-1999
DEFINITION Rattus norvegicus mRNA for N-G,N-G-dimethylarginine
            dimethylaminohydrolase, complete cds.
ACCESSION   D86041
VERSION   D86041.1 GI:1906799
KEYWORDS   N-G,N-G-dimethylarginine dimethylaminohydrolase.
SOURCE   Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (sites)
AUTHORS   Kimoto,M., Sasakawa,T., Tsuji,H., Miyatake,S., Oka,T., Nio,N. and
            Ogawa,T.
TITLE   Cloning and sequencing of cDNA encoding NG,NG-dimethylarginine
            dimethylaminohydrolase from rat kidney
JOURNAL   Biochim. Biophys. Acta 1337 (1), 6-10 (1997)
MEDLINE   97157050
PUBMED   9003431
REFERENCE   2 (bases 1 to 3008)
AUTHORS   Kimoto,M.
TITLE   Cloning and sequencing of cDNA encoding G-N, G-N-dimethylarginine
            dimethylaminohydrolase from rat kidney
JOURNAL   Unpublished
REFERENCE   3 (bases 1 to 3008)
AUTHORS   Kimoto,M.
TITLE   Direct Submission
JOURNAL   Submitted (14-JUN-1996) Masumi Kimoto, The University of Tokushima,
            School of Medicine, Nutrition; Kuramoto 3, Tokushima, Tokushima
            770, Japan (E-mail:kimoto@nutr.med.tokushima-u.ac.jp,
            Tel:0886-33-7088,Fax:0886-33-7089)
FEATURES   Location/Qualifiers
     source   1..3008
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /sex="female"
                     /tissue type="kidney"
     CDS   432..1289
                     /note="DDAH"
                     /codon_start=1
                     /product="N-G,N-G-dimethylarginine dimethylaminohydrolase"
                     /protein_id="BAA18993.1"
                     /db_xref="GI:1906800"
                     /translation="MAGLSHPSVFGRATHAVVRAPPESLCRHALRRSGEEVDFARAE
                     RQHQLYVGVLGSKLGLQVVQLPADESLPDCVFVEDVAVVCEETALITRPGASRRKEV
                     DMVKEALEKLQLNIVEMKDENATLDGGDVLFTGREFFVGLSKRTNQRGAEILADTFKD
                     YAVSTVPVADSLHLKSFCSMAGPNLIAIGSSESAQKAIKIMQQMSDHRYDKLTVPDDM
                     AANCIYLNIPSKGHVLLHRTPEEYPESAKVYEKLKDHLLIPVSNSEMEKVDGLLTCCS
                     VFINKKTDS"
ORIGIN

Query Match   83.4%;   Score 715.6; DB 10; Length 3008;
Best Local Similarity   89.6%;   Pred. No. 2.4e-108;
Matches   769; Conservative   0; Mismatches   89; Indels   0; Gaps   0;

Qy   1 ATGGCCGGCCTCGGCCACCCCTCGGCCCTTCGGCCGGCCACCCAGCCGTTGGTGCGGGCG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   432 ATGGCCGGCCTCAGCCACCCCTCGGTCTTCGGCCGGCCACCCAGCCGTTGGTGCGGGGCT   491

Qy   61 CTACCCGAGTCGCTCTGCCAGCACGGCAGCGCCAGGGCCGCCAAGGGCGGAGGTGGACGTC 120
          |||||||||||||| ||||| ||||||||| ||| |||| | ||||| || ||||||||||
Db   432 ATGGCCGGCCTCAGCCACCCCTCGGTCTTCGGCGGGCCACCCAGCCGTTGGTGCGGGGCT   491

Qy   61 CTACCCGAGTCGCTCTGCCAGCACGGCAGCGCCAGGGCCGCCAAGGGCGGAGGTGGACGTC 120

Qy   432 ATGGCCGGCCTCAGCCACCCCTCGGTCTTCGGCGGGCCACCCAGCCGTTGGTGCGGGGCT   491

TITLE   Imaki,J., Suzuki,H. and Yamashita,K.
JOURNAL   Direct Submission
           Submitted (30-OCT-2002) Takuya Mishima, Nippon Medical School,
           Department of Anatomy; Sendagi 1-1-5, Bunkyo, Tokyo 113-8602, Japan
           (E-mail:mishima@nms.ac.jp, Tel:81-3-3822-2131(ex.5315))
COMMENT   On Jun 26, 2003 this sequence version replaced gi:30268647.
FEATURES   Location/Qualifiers
     source   1..1150
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9031"
                     /tissue_type="embryos yolk sac"
                     /clone_lib="Chick Yolk Sac Lambda cDNA library (Stratagene
                     #937406)"
                     /dev_stage="Five-days-old chick"
     CDS   40..903
                     /codon_start=1
                     /product="dimethylarginine dimethylaminohydrolase I"
                     /protein_id="BAC76002.1"
                     /db_xref="GI:3026864B"
                     /translation="MAGLGGPAAFGRCTHAVVRALPESLCRQALRSTAGEVDFARA
                     BREHQLYVGVLRGKLGLQVLEIPADESLPDCVFVEDAAVVCBETALLTRPGAPSRRKE
                     VEAMKRVLESLNLNMVEMVDENATLDGGDVLFTGREFFFVGLSKRTNQRGAEILADTF
                     KDYAVSTVPVHDSLHLKSFCSMAGPNLIAIGSSEAQKALKTMQQWSDHRYDKLTVPD
                     DAAANCIYLNIPSKGHVLLHRAPEEYPESAKVFEKLKDHMLIPIANTELEKVDGALTC
                     CSVLINKTSEL"

ORIGIN
Query Match   67.6%; Score 579.6; DB 5; Length 1150;
Best Local Similarity   80.9%; Pred. No. 6.4e-86;
Matches 688; Conservative 0; Mismatches 159; Indels 3; Gaps 1;

Qy   12 CGGGCCACCCTCCGCCTTCGGCCGGCCACCACGCCCGGTGGTGGGGCTACCGCGAGTC 71
          |||||||||||||||||||| |||||| |||| ||| |||| |||||||||||||
Db   54 CGGGCCACCCTCCGCCTTCGGCCGGCCGCCACGCCCGGTGGTGGGGCTGCCGCCGAGTC 113

Qy   72 GCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGCCGAGGAGCGTCGCGCCGCGGA 131
          ||||| ||||||||||||| ||||| ||||| |||||| ||||||||||||||||
Db   114 GCTGTGCCGGCACGGCGTTGCGAGCACGCGGCCCCGAGGTGACTTCGCTCGCGCCGA 173

Qy   132 ACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAAGCTGGGCGTGCAGGTGGTGTGA 191
          |||| |||| |||| |||||| |||| |||| |||| ||||| |||||||||||||||
Db   174 GCGGGAGCATCAGCTGTACGTGGGCGTGCTACGGCAAGCTGGGCGTGCAGGTGCTGA 233

Qy   192 GCTGCCGGCCGACGAGAGCCTTCGGACGCGTCTTCGTGGAGCGTGCCCTGGCCGTGTG 251
          |||| |||||||||||| |||||| ||||| |||||||| |||||||||| |||||||
Db   234 GCTGCCGGCTGACGAGAGCCTCCCGACGCGTCGTCGTGGAGCGTCGCGCCGGCCGTGTG 293

Qy   252 CGAGGAGACGGCCCTCATCACCCGACCCGGGGGCGCCGAGGAGGAGGTTGACAT 311
          |||||||||||||||||| ||||| ||||||||||||||| ||||||||||||
Db   294 CGAGGAGACGGCGCTGCTCACCCCGCCCCCCAGCCGCGCCGAGGAGGAGGTTGAAGC 353

Qy   312 GATGAAAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGCAGCAGAGAA 371
          |||||| |||||||| ||||| ||| |||| ||||| ||| |||
Db   354 CATGAAGAGAGTACTAGAAAAGCTTAACCTCAACATGGTAGATGGTAGAGATGAAAAATGC 413

Qy   372 AACTTTAGATGGGGAGATGTTTATTCACAGGCAGAGAA----TTTTTGTGGGCCTTTC 428
          ||||||||| |||||||| ||||| ||||| ||||| | |||||||| |||||||
Db   414 AACCTTGGATGGTGGAGACGTCTTATTTACAGGCAGAATTTTTTTTTGTAGGCCTTTC 473

Qy   429 CAAAAGGACAAAATCAACGAGAGTGCTGAAATCTTGGCTGACACTTTTAAGGACTATGCAGT 488
          ||||| |||||||||||| |||||||| |||| |||||||| ||||||||||| ||||||
Db   474 CAAGGCGGACAAATCAACCAACTGGGCCAGAAATTCTGGCTGCACAATTTAGGATTATGTGT 533

Qy   489 CTCCAGTGCCAGTGGCAGATGGTTGCATTGAAGAGTTTCTGCAGCATGGTGGGCC 548
          |||| ||| |||||||| |||||| |||| ||||||||||||| |||||||||||
Db   534 CTCCACAGTCCCGTCGTTCATGATTCTTGCATCTGAAGAGTTTTGCAGCATGGCTGGACC 593

Qy   549 TAACCTGATCGCAATTGGTCTAGTGAATCTGCACAGGAGGCCCTTAAGATCATGCAACA 608
          |||||||||||||||| |||||||||||||||||||||||||| |||||||||||||||
Db   594 AAACCTGATTGCTATCGGGTCAAGTGAAGCTGCACAGAAAGCCCTCAAGACCATGCAACA 653

Qy   609 GATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACTACGAGCAGCAAACTGTAT 668

Qy   121 GCCCGCGCGGAACGGCAGCAGCCTCTACGTGGGCGTGCTGGGCAAGCTGGGGGCTG 180
          |||| ||||||||||||||||||||||| |||||||| ||||||||||||| |||
Db   552 GCTCCGGCGCTAGCGCGCCAGCACCAGTCTCACGTGGGCGTCGTGCTGGGCCAGCAAGCTGGGGGTG 611

Qy   181 CAGGTGGTGGAGCTGCCGCCGGCCACCAGAGCCTTCGGACTGCGTCTTCGTGGAGGACGTG 240
          ||||||||||| |||||||||||||||||||| |||||| |||||||||||||||||||
Db   612 CAGGTGGTGCAGCTGCCGCCGCCACAGAGAGCCTTCGGACTGCGTCTTCGTGGAGGACGTG 671

Qy   241 GCGGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGGCGCCGAGGAGCGGAAG 300
          ||||||||| ||||||||||| ||||| ||||| ||||||||||||||||||||||||
Db   672 GCGGTGGTGTGCGAGGAGACGGCGCTGCTCACCCCGCCCCCCAGCCGCGCCGAGGAGCGGAAG 731

Qy   301 GAGGTTGACATGATGATGAAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGATGAGAA 360
          |||||||| |||||||||| ||||||| ||||||||| |||| |||||||||||||
Db   732 GAGGTTGACATGATGATGAAAGAGAGGAGCTCCTATTCACAGGCAGAGAGTTTTTTGTG 791

Qy   361 GATGAAAATGCAACTTAGATGGCCGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 420
          ||||||||||||||||||||| ||||||||||||| ||||| ||||| ||||||||||
Db   792 GATGAAAATGCAACCTTAGATGGTGGGGACTCCTCATTCACAGGCAGAGAGTTTTTGTG 851

Qy   421 GGCCTTTCCAAAAGGACAAATCAACGAGAGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
          ||||||||||||| |||||||||||| |||||| |||||||||||||| |||||||||||
Db   852 GGCCTTTCCAAAAGGACAAATCAACGAGTCTGCTGAGATCTTGGCTGATACTTTTCAAGGAC 911

Qy   481 TATGCAGTCTCCACAGTGGTTGCAGTGGTTGCATTGAAGAGTTTCTGCAGCATG 540
          |||| |||||||| |||||||||||| |||||||||||||||||||||||||||
Db   912 TACGCAGTTTCCACAGTCCCCGTGGCCGATTCTTGCATTTGCATTTAAAGAGTTTCTGCAGCATG 971

Qy   541 GCTGGGCCTAACCTGATCGCAATTGGTCTCAGTGAATCTGCACAGAAGGCCCTTAAGATC 600
          |||||| ||||||||||||||| |||||| ||||| |||||||||| |||||||||||||
Db   972 GCTGCCCCAACCTGATCGCAATAGGGTCTCCAGTGAATCTGCGCAGGAGGCCCTCAAGATC 1031

Qy   601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACTAGCAGCA 660
          |||||||||||| ||| |||||||| ||||||||||||| |||| || || |||||
Db   1032 ATGCAACAGATGAGTGACCACCGTTATGACAAGCTCACTGTTCACCGACACATGCCGCC 1091

Qy   661 AACTGTATATATCTAAATATCCCAACAAAAGGGCACGTCTTGCTGCACCGGAACCCCGGAA 720
          |||||||| || | ||||||||| ||||||||||||| |||||| |||||||||||||||
Db   1092 AACTGTATATTTAAATATCCCCAGCAACAAGGGCATGTCTTGCTGCTCCACCGGAACCCCAGAA 1151

Qy   721 GAGTATCCAGAAAGTGCAAAGTTTATGAAGAACTGAAGGACCATGTGTCATCCCCGTG 780
          ||||||||||| ||||||| |||||||||| ||| |||| ||| |||| ||||||||
Db   1152 GAGTACCCAGAAAGCCGCAAGGTTTATGAAGGTTCAAGAGCCATCTCACTACTGATCCCTGTG 1211

Qy   781 AGCATGTCTGACTGGAAAAGGTGGATGGCTGCTCACCTGCTCAGTTTAATTAAC 840
          |||||| ||||||||||||||||| || |||||||||| |||||| |||||||||
Db   1212 AGCAATTCTGAGATGGAAAAGGTGGACGGCTGCTCACCTGCTCGTGCTGTTTTTATTAAC 1271

Qy   841 AAGAAGTAGACTCCTGA 858
          |||||||||| ||||||
Db   1272 AAGAAGAGACAGACTCTGA 1289

RESULT 11
AB095027
LOCUS   AB095027   1150 bp   mRNA   linear   VRT 27-JUN-2003
DEFINITION Gallus gallus mRNA for dimethylarginine dimethylaminohydrolase I,
            complete cds.
ACCESSION   AB095027
VERSION   AB095027.2 GI:32306458
KEYWORDS   .
SOURCE   Gallus gallus (chicken)
  ORGANISM Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1
  AUTHORS   Mishima,T., Hamada,T., Ui-Tei,K., Takahashi,F., Miyata,Y.,
            Imaki,J., Suzuki,H. and Yamashita,K.
  TITLE   The expression of DDAH I in the chick and rat embryo
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1150)
  AUTHORS   Mishima,T., Hamada,T., Ui-Tei,K., Takahashi,F., Miyata,Y.,

Left column:

```
Db  654 GATGAGCGACCACCGCTACGACCAAGCTGACGGTGCCCGACGATGCGCCGCCGAAACTGCAT 713
            ||||||  ||||  || |||    ||||    || || |||  ||||||||
Qy  669 ATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCCACCGGACCCCGGAAGAGTATCC  728
            || ||| |||||||||||    ||||||||   ||||||  |||||||| |||||||
Db  714 CTACTTAAACATTCCCAGCAAAGGGCACGTCCACCCAGCCCCTGAGGAGTACCC 773
Qy  729 AGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTC 788
            ||| || |||||||||  |||||||||||||||||| |||||||| |||  ||  |
Db  774 AGAGAGCGCAAAGGTTTTTGAAAAACTGAAGGACCACATGCTGATCCCATAGCCAACAC 833
Qy  789 TGAACTGGAAAAGGTGGATGGGCTGCTGCTCACCTGCTCGTCCCAGTTTAATTAACAAGAAGGT 848
            |||||||| |||| |||||||  ||  |||||||| |||| || |||||||| |||| |
Db  834 AGAACTGGAGAAGTAGATGGGGGCACTCACCTGTTGCTCTGTTATTAACAAAAACTTC 893
Qy  849 AGACTCCTGA 858
            ||   | ||
Db  894 AGAATTATGA 903
```

```
RESULT 12
HSM808292
LOCUS       HSM808292   3908 bp   mRNA   linear   PRI 30-AUG-2003
DEFINITION  Homo sapiens mRNA; cDNA DKFZp686N2176 (from clone DKFZp686N2176).
ACCESSION   BX648145
VERSION     BX648145.1  GI:34367304
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3908)
  AUTHORS   Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  CONSRTM   The German Human cDNA Consortium
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            This clone (DKFZp686N2176) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1..3908
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686N2176"
                     /tissue_type="686 (synonym: hlcc3). Vector pSport1_Sfi; host
                     DH10B; sites SfiIA + SfiIB"
                     /clone_lib="human endometrium carcinoma cell line"
                     /dev_stage="adult"
ORIGIN

Query Match      64.8%;  Score 555.4;  DB 9;  Length 3908;
Best Local Similarity  99.8%;  Pred. No. 6.1e-82;
Matches  556;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;

Qy  302 AGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCCAATATAGTAGAGATGAAAG 361
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  382 AGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCCAATATAGTAGAGATGAAAG 441
Qy  362 ATGAAAATGCAACTTTAGATGGGCGGAGATGTTTATTCACAGGCAGCAGAGAATTTTTGTGG 421
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  442 ATGAAAATGCAACTTTAGATGGGCGGAGATGTTTATTCACAGGCAGCAGAGAATTTTTGTGG 501
```

Right column:

```
Qy  422 GCCTTTCCAAAAGGACAAATCAAGGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACT 481
            ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  502 GCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACT 561
Qy  482 ATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTCTGCAGCATGG 541
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  562 ATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTCTGCAGCATGG 621
Qy  542 CTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAGAGGCCCTTAAGATCA 601
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  622 CTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAGAGGCCCTTAAGATCA 681
Qy  602 TGCAACAGATGAGTGACCACCGCTACGACAAAACTCACTGTGCCTGATGACATAGCAGCAA 661
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  682 TGCAACAGATGAGTGACCACCGCTACGACAAAACTCACTGTGCCTGATGACATAGCAGCAA 741
Qy  662 ACTGTATATATCTAAATATCCCAACAAAGGCAGTCTTGCTGCCACCGGACCCCGGAAG 721
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  742 ACTGTATATATCTAAATATCCCAACAAAGGCAGTCTTGCTGCCACCGGACCCCGGAAG 801
Qy  722 AGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGA 781
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  802 AGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGA 861
Qy  782 GCATGTCTGAACTGGAAAAGGTGGATGGGCTTGCTCACCTGCTCCAGTTTAATTAACA 841
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  862 GCATGTCTGAACTGGAAAAGGTGGATGGGCTTGCTCACCTGCTCCAGTTTAATTAACA 921
Qy  842 AGAAGTAGACTCCTGA 858
            |||| ||||||||||
Db  922 AGAAAGTAGACTCCTGA 938
```

```
RESULT 13
AK128909
LOCUS       AK128909   2289 bp   mRNA   linear   ROD 09-SEP-2003
DEFINITION  Mus musculus cDNA fis, clone TRACH2023479, highly similar to
            NG,NG-DIMETHYLARGININE DIMETHYLAMINOHYDROLASE 1 (EC 3.5.3.18).
ACCESSION   AK128909
VERSION     AK128909.1  GI:34536566
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
            Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
            Kamihara,H., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
            Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
            Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y.,
            Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2289)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO cDNA sequencing project supported by Ministry of Economy,
            Trade and Industry of Japan; cDNA full insert sequencing: Research
            Association for Biotechnology (RAB); cDNA library construction:
            Helix Research Institute (HRI) (supported by Japan Key Technology
            Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and
            Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1..2289
```

```
                    /organism="Mus musculus"
                    /mol_type="mRNA"
                    /db_xref="taxon:10090"
                    /clone="TRACH202479"
ORIGIN              /note="cloning vector: pME18SFL3"

Query Match    53.9%;  Score 462.4;  DB 10;  Length 2289;
Best Local Similarity 89.1%;  Pred. No. 1.3e-66;
Matches 499;  Conservative  0;  Mismatches  61;  Indels  0;  Gaps  0;

Qy    299  AGGAGGTTGACATGATGAAAGAAGCATTAGAAAAAACTTCAGCTCAATATAGTAGAGATGA  358
                |||||||||||||||||| || ||||| || |||||||||||||||||||||| ||||||||
Db     24  AGTAGGTTGACATGATGAAAGAAGAGGCTTTGGAAAAAACTTCAGCTCAACATAGTAGAGATGA  83

Qy    359  AAGATGAAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTG  418
                |||||||||||||||||||||||||||  ||||| | |||| |||||||||||||||||
Db     84  AAGATGAAAATGCAACTTTCGGATGTGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTTG  143

Qy    419  TGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGG  478
                ||||||||||||||||||| ||||||||||| ||||| || ||||||| |||||||||||
Db    144  TGGGCCTTTCCAAAGAACAAATCAACGAGGTGCTGAAAATCTTGGCTGATACTTTTAAGG  203

Qy    479  ACTATGCAGTCTCCACAGTGCCAGTGCGAGATGGGTTGCATTGAAGAGTTTCTGCAGCA  538
                ||||||||||||| |||||||  |||||||||||||||||| |||||||||||||||||
Db    204  ACTACGCAGTCTCTACAGTCCCTGTGGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCA  263

Qy    539  TGGCTGGCCTAACCTGATCGCCAATTGGGTCTAGTGAATCTGACACAGAAGGGCCCTTAAGA  598
                ||||||  |||||||||||||||||||| |||||| ||||||||||||||||||||||||
Db    264  TGGCCGGACCCAACCTGATTGCAATAGGGTCCAGCAAATCTGACAGAAGAGGGCCCTCAAGA  323

Qy    599  TCATGCAACAGATGAGTGACCACCGGCTAGACAAACTCACTGTGCCTGATGACATAGCAG  658
                |||||||||||||||| |||||||||| |||||||||||||| |||||||||| |||||
Db    324  TCATGCAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGACATGGGCCG  383

Qy    659  CAAACTGTATATATCTAAATATCCCCAACAAAGGGCACGTCCTTGCTGCACGGACCCCGG  718
                |||||||||||||| |||||||| |||||||||||||||||||||||||| |||| ||||
Db    384  CCAACTGCATATATCTAAATATCCCCAGCAAAGGGCCAAAGGCCATGTCTTGCTGCACCGGACCCCAG  443

Qy    719  AAGAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCG  778
                ||||| ||||||||| ||||||||| |||||||||||||||| |||| ||||||||||||
Db    444  AAGAGTACCCAGAAAGCGCAAAGGTCGTATAGAAACTCAAGGACCATCTACTGATCCCTG  503

Qy    779  TGAGCATGTCGAACTGGAAAGGTGGAAAGGTGGATGGGCTGCTCACCTGCTCCAGTTTTAATTA  838
                |||||| |||||||||||||||  |||||||||||||||||||||||||||||  ||||||
Db    504  TGAGCAACTCGGAGATGGAAAGGTGGACGGGCTTGCTCACCTGCTCCCGTTTTTTATTA  563

Qy    839  ACAAGAAGTAGACTCCTGA  858
                ||||||||| |||||||||
Db    564  ACAAGAAGATAGACTCCTGA  583
```

RESULT 14
BC056074
LOCUS   BC056074   1344 bp   mRNA   linear   VRT 04-NOV-2003
DEFINITION Xenopus laevis dimethylarginine dimethylaminohydrolase 1, mRNA
            (cDNA clone MGC:69055 IMAGE:4965611), complete cds.
ACCESSION   BC056074
VERSION   BC056074.1 GI:33417137
KEYWORDS   MGC.
SOURCE   Xenopus laevis (African clawed frog)
ORGANISM   Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1 (bases 1 to 1344)
AUTHORS   Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
            and Richardson,P.
TITLE   Genetic and genomic tools for Xenopus research: The NIH Xenopus
            initiative
JOURNAL   Dev. Dyn. 225 (4), 384-391 (2002)
MEDLINE   22341132
PUBMED   12454917

REFERENCE   2 (bases 1 to 1344)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22388257
PUBMED   12477932
REFERENCE   3 (bases 1 to 1344)
AUTHORS   Klein,S. and Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (01-AUG-2003) National Institutes of Health, Xenopus Gene
            Collection (XGC), National Institute of Child Health and Human
            Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
            20892-7510, USA
REMARK   NIH-MGC Project
COMMENT   Contact: XGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Igor Dawid
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site: http://www-shgc.stanford.edu
            Contact: (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R.M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 129 Row: J Column: 20
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein.
FEATURES   Location/Qualifiers
     source   1..1344
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="MGC:69055 IMAGE:4965611"
                     /tissue_type="Spleen, adult Xenopus"
                     /clone_lib="NICHD_XGC_Spl"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene   1..1344
                     /gene="ddah1-prov"
                     /note="synonym: MGC69055"
                     /db_xref="LocusID:380577"
     CDS   84..938
                     /codon_start=1
                     /product="ddah1-prov protein"
                     /protein_id="AAH56074.1"
                     /db_xref="GI:33417138"
                     /db_xref="LocusID:380577"
                     /translation="MAGLALSGFGRCTHVIVRAIPDSLCQEALRIEKGSVVDISLARK
                     QHEQYYEVLKNKLGLSVIELPADESLPDAPFVEDVAVVCEETALITRPGAPSRRKETE
                     AVKKALQDLQLNVEWTDENATLDGGDVLFTGREFFVGLSKRTNQRGAEILADTFKDY
```

```
882 GTTGATGGATGTCTGACCTGCTGCTCCATCCTTATTAACAAAAAGAAAAGTCTCTGA 938
```

Db

RESULT 15

BC060266

LOCUS   BC060266   3420 bp   mRNA   linear   ROD 25-NOV-2003
DEFINITION Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
            clone IMAGE:5698949), complete cds.
ACCESSION   BC060266
VERSION   BC060266.1 GI:37805213
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 3420)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22388257
PUBMED   12477932
REFERENCE   2 (bases 1 to 3420)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (20-OCT-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 132 Row: a Column: 19
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis
            This clone has the following problem: The cds is short compared to
            the longest cds in the locus.
FEATURES   Location/Qualifiers
            misc_feature
```

```
            AVSTVPVCETHLKSFCSMAGPNLIAIGSSEAAQKALKTMQQMSDHRYDKLTLPDNAA
            ANCIYLNIPSKGHVLLHLAPEDFPESAKVFEKLKEHMLFPLPNTELAKVDGLITCCSI
            LINKKKSL"
            114..917
            /note="Amidinotransf; Region: Amidinotransferase. This
            family contains glycine (EC:2.1.4.1) and inosamine
            (EC:2.1.4.2) amidinotransferases, enzymes involved in
            creatine and streptomycin biosynthesis respectively. This
            family also includes arginine deiminases, EC:3.5.3.6.
            These enzymes catalyse the reaction: arginine + H2O <=>
            citrulline + NH3. Also found in this family is the
            Streptococcus anti tumour glycoprotein"
            /db_xref="CDD:pfam02274"
```

ORIGIN

Query Match   53.4%;   Score 457.8;   DB 5;   Length 1344;
Best Local Similarity 71.7%;   Pred. No. 7.7e-66;
Matches 600;   Conservative 0;   Mismatches 237;   Indels 0;   Gaps 0;

```
Qy    22 TCCGCCTTCGGCCGGGCCACCGCCGTGGTGCGGGGCGTACCGAGTCGCTCTGCCAG  81
          ||||| || | ||| |||||| |||| ||||| || || || ||||| |||||||
Db   102 TCAGGCTTTGGGAGATGCACCCACGTTATAGTTCGTGCAATTCCTGACTCTTTATGTCAA 161

Qy    82 CACCGCTGAGAGAGCGCCAAGGGCGAGGAGGTGACGTCGCCCGCGGAACGGCAGCAC 141
          || | || || ||| ||||| |||| ||||| || || ||| ||| || |||||
Db   162 GAAGCTCTGAGGATTAGAGAAGGCAGTGTGGTGGTGATATCCCCTGGCTAGGAAACAACAT 221

Qy   142 CAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCTTGCAGGTGGTGGAGCTGCCGCC 201
          |||| |||||| | ||||| | ||| | ||||| ||||| ||||| || | ||| |||
Db   222 GAGCAGTATGTTGAGGTGTCTGAAGAATAAACTGGGGCTGAGTGTGATAGAGCTTCCAGCA 281

Qy   202 GAGCAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTCGGAGGACG 261
          |||| || |||| || || ||||| | ||| | || ||||| ||||||| |||| |||
Db   282 GATGAATCCCTCCCAGATGCGCCCTTTGTGGAAGATGTGGCTGTTGTGTGTGAGGAAACT 341

Qy   262 GCCCTCATCACCCGACCCGGGCGCCCGAGCCGGAGGAAGGAGGTTGACATGATGAAAGAA 321
          ||| || ||| || ||||||||| || |||| | || ||||| ||||| | | || || |
Db   342 GCCCGATTACCCAGACCTGGGGGGCACCTAGCCAGGAGGAAAGAGACTGAAGCTGTGAAGAAA 401

Qy   322 GCATTAGAAAAACTTCAGCTCAATATAGTAGATGAGATGAAAATGCAACTTAGAT 381
          ||| |||||||| || ||||| | || || ||||| | || ||| ||| |||||||
Db   402 GCACTGCAGGATCTTCAGCTAAATGTGGTTGAAATGACAGATGAAAATGCTACATTAGAT 461

Qy   382 GGCGGAGATGTTTTATTCACAGGACAGAGAATTTTTGTGGCCTTTCCAAAAGGACAAAT 441
          ||||| ||||||||  |||| ||| ||||  |||| ||| ||| ||||| | || |||
Db   462 GGAGGAGATGTTTTTGTTTACAGGAAGGGAATTCTTTGTGGGTTTGTCGAAAAGGACCCAAT 521

Qy   442 CAACGAGGTGCTGAAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCA 501
          ||| || |||||||| | ||||||| || || ||||| || || |||| ||||| |||
Db   522 CAACGTGGTGTCTGAAAATTCTGGCAGATACCTTTTAAGGATTATGCAGTTTTCCACTGTGCCA 581

Qy   502 GTGTGCAGATGGGTTGCATTGAAGAGTTTGAAGGGCTCGGGCTTAACCTGATCGCA 561
          ||| | ||| |||| || |||| | ||||| ||| ||| |||| |||| | || ||
Db   582 GTGTGCGGAGCAATTGCACTTGAAGAGTTTCTGCGAGCATGGCTCGGAGCCAAACCTCATTGCC 641

Qy   562 ATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCAC 621
          ||| ||||| | || | |||| |||| |||| ||| | || ||| |||| || ||||||
Db   642 CCAAGCAAAGGGCATGTTCTACTGCATCTTGCACCTGACATGGGCATGTCTGAGCATGGAAAAG 701

Qy   622 CGCTACGACAAATCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATC 681
          || || ||||| ||| ||| || || |||| |||| ||| |||| | ||  ||||| |
Db   702 CGCTACGACAAGCTGACAATTGCCTGATAATGCTGCAGCTAACTGCATTTATTTAAACATT 761

Qy   682 CCAACAAAGGGCACGTCTTGTCTGCACCGAACCCGGAAGATATCCGAGAAGTGCAAAG 741
          || ||||| ||| | |||| ||| || |||||| |||| |||  ||| | ||| |||
Db   762 CCAAGCAAAGGGCATGTTCTACTGCATCTTGCACCTGACATGTTTCAAGAGAGTGCAAAG 821

Qy   742 GTTTATGAGAACTCAAGGAACCATATGCTAATCCCGTGAGCATGTCTGAACTGGAAAAG 801
          |||| || ||||| || || | |||||| || ||| | || |||  ||| |||||||
Db   822 GTGTTTGAAAACTAAAGGAACCACTGCGTGTTTCCACTTCCTAACACCGAGCTTGCAAA 881

Qy   802 GTGGATGGGCGCTGCTCACCTGCTGCTCAGTTTAATTAACAAGAAGGTAGACTCCTGA 858
          ||| ||||| | ||  |||  ||
```

Search completed: June 7, 2004, 20:25:47
Job time : 3417 secs

```
source          1..3420
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6"
                /db_xref="taxon:10090"
                /clone="IMAGE:5698949"
                /tissue_type="Brain, whole, 18.5 dpc mouse embryo"
                /clone_lib="NIH_BMAP_EQ0"
                /lab_host="DH10B"
                /note="Vector: pYX-ASC"

gene            1..3420
                /gene="Ddah1"
                /note="synonym: 2510015N06Rik"
                /db_xref="LocusID:69219"
                /db_xref="MGI:1916469"

CDS             360..680
                /codon_start=1
                /product="Ddah1 protein"
                /protein_id="AAH60266.1"
                /db_xref="GI:37805214"
                /db_xref="LocusID:69219"
                /translation="MAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDMAANCI
                YLNIPSKGHVLLHRTPEEYPESAKVYEKLKDHLLIPVSNSEMEKVDGLLTCCSVFINK
                KIDS"

misc_feature    360..659
                /note="Amidinotransf; Region: Amidinotransferase. This
                family contains glycine (EC:2.1.4.1) and inosamine
                (EC:2.1.4.2) amidinotransferases, enzymes involved in
                creatine and streptomycin biosynthesis respectively. This
                family also includes arginine deiminases, EC:3.5.3.6.
                These enzymes catalyse the reaction: arginine + H2O <=>
                citrulline + NH3. Also found in this family is the
                streptococcus anti tumour glycoprotein"
                /db_xref="CDD:pfam02274"

ORIGIN

Query Match         44.5%;   Score 381.8;  DB 10;  Length 3420;
Best Local Similarity   88.8%;   Pred. No. 2.5e-53;
Matches   413;   Conservative    0;   Mismatches   52;   Indels    0;   Gaps    0;

Qy    394 TTATTCACAGGCAGAGAATTTTTGTGGGCCTTCCAAAAGGACAAATCAACGAGGTGCT 453
          ||||| |||||||||||||||||||||||||| |||||||||| ||||||||||||||
Db    216 TTGCTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGAACAAATCAACGAGGTGCT 275

Qy    454 GAAATCTTGGCTGATACTTTTAAGGACTATTGCAGTCTCCACAGTGCCAGTGCAGATGGG 513
          |||||||||||||||||||| ||||||||| |||||||||| |||||||||||||| ||
Db    276 GAAATCTTGGCTGATACTTTTAAGGACTACGCAGTCTCTACAGTCCCTGGCCGATTCT 335

Qy    514 TTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGT 573
          |||||| |||||||||||||||||||||||| |||||| |||| |||||||||||||||
Db    336 TTGCATTTAAAGAGTTTCTGCAGCATGGCCGGGCCGACCCAACCTGATTGCAATAGGGTCCAGC 395

Qy    574 GAATCTGCACAGAAGGCCCCTTAAGATCATGCAACAGATGAGTGAGTGCCACCGCTACGACAAA 633
          ||||||||||||||||||||||||||| ||||||||||| |||||| |||||||||||||||
Db    396 GAATCTGCACAGAAGGCCCCTCAAGATCATGCAACAGATGAGTGAGTGCCATCGTTATGACAAG 455

Qy    634 CTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCAACACTGTATATATCCCAACAAAGGG 693
          |||||||||||||||||| ||||||||| |||||||||| |||||||||||| |||||||||||
Db    456 CTCACTGTACCCGACGACATGGCCGCCGCCAACTGCATATATCCAACACATATATCCCAGCAAAGGG 515

Qy    694 CACGTCTTGCTGCACCGAACCCGGAAGAGTATCCGAAGTGCAAAGTTTATGAGAAA 753
          ||||||| |||||||||||||| |||||||||||||| |||||||||| |||||||
Db    516 CATGTCTTGCTGCACCGAACCCGTGAGCAACTGGAAGAGCGCAAAGGTCTATGAGAAA 575

Qy    754 CTGAAGGACCATATGCTGAGATCCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTG 813
          |||||||||||||||| |||| ||||| |||||| |||||||||||||| |||| ||||||
Db    576 CTCAAGGACCATCTACTGATCCCTGTGAGCAACTCGGAGATGGAAAAGGTGGACGGGCTTG 635

Qy    814 CTCACCTGCTGCTCAGTTTAATTAACAAGAAGGTAGACTCCTGA 858
          |||||||||||||| ||||| ||||||||||| ||||||||||
Db    636 CTCACCTGCTGCTCCGTTTTTATTAACAAGAAGATAGACCCCTGA 680
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2004, 22:28:56 ; Search time 3208 Seconds
(without alignments)
3850.612 Million cell updates/sec

Title: US-09-889-733B-2
Perfect score: 1454
Sequence: 1 MAGLGHPSAFGRATHAVVRA........EKVDGLLTCCSVLINKKVDS 285

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
MODEL=frame+_p2n.model -DEV=xlp
Q=/cgn2.1/USPRO_spool_p/US09889733/runat_07062004_083528.7970/app_query.fasta_1.455
DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US09889733 @CGN_1_1_5600 @runat_07062004_083528.7970 -NCPU=6 -ICPU=3
NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=30 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1454	100.0	858	6	BD251986	BD251986 Screen me
2	1454	100.0	858	6	AX032835	AX032835 Sequence
3	1439	99.0	1624	6	BC033680	BC033680 Homo sapi
4	1439	99.0	1633	6	BD094073	BD094073 Shear str
5	1439	99.0	1633	9	AB001915	AB001915 Homo sapi
6	1439	99.0	4208	9	AX548046	AX548046 Sequence
7	1389	95.5	1120	10	BC034505	BC034505 Mus muscu
8	1379	94.8	3008	6	AX683109	AX683109 Sequence
9	1379	94.8	3008	10	D86041	D86041 Rattus norv
10	1253	86.2	3737	9	BC043235	BC043235 Homo sapi
11	1243	85.5	1150	5	AB095027	AB095027 Gallus ga
12	1149.5	79.1	1344	9	BC056074	BC056074 Xenopus l
13	941	64.7	3908	9	HSM808292	BX648145 Homo sapi
14	910	62.6	2289	10	AK128909	AK128909 Mus muscu
15	756	52.0	3420	10	BC060266	BC060266 Mus muscu
16	696	47.9	1310	10	BC003328	BC003328 Mus muscu
17	694	47.7	1263	9	BC001435	BC001435 Homo sapi
18	694	47.7	1342	9	AK098671	AK098671 Homo sapi
19	694	47.7	1351	6	AX779950	AX779950 Sequence
20	694	47.7	1351	9	AF087894	AF087894 Homo sapi
21	694	47.7	1376	9	AX011712	AX011712 Sequence
22	694	47.7	1376	6	BD226323	BD226323 Pancreati
23	694	47.7	1709	9	AK026191	AK026191 Homo sapi
24	688	47.3	858	6	BD251987	BD251987 Screen me
25	688	47.3	858	6	AX032837	AX032837 Sequence
26	661.5	45.5	1270	10	AF004106,	AF004106, Mus muscu
27	645.5	44.4	1397	9	AR379492	AR379492 Sequence
28	624	42.9	1228	9	AX780306	AX780306 Sequence
29	546	37.6	2895	9	AK098405	AK098405 Homo sapi
30 c	502	34.5	104071	2	AL162260	AL162260 Homo sapi
31	502	34.5	158405	2	AC019261	AC019261 Homo sapi
32 c	502	34.5	168953	2	AL360219	AL360219 Human DNA
33	483	33.2	186434	2	AC126412	AC126412 Mus muscu
34	482	33.1	1377	9	AY060959	AY060959 Drosophil
35 c	476	32.7	228940	2	AC126890	AC126890 Rattus no
36	474	32.6	135545	10	MMHC213L3	AF109905 Mus muscu
37	474	32.6	221899	10	AC087117	AC087117 Mus Muscu
38	449.5	30.9	228453	2	AC094348	AC094348 Rattus no
39	423	29.1	28772	9	HSA012008	AJ012008 Homo sapi
40 c	423	29.1	53125	9	AL844216	AL844216 Human DNA
41	423	29.1	54146	9	BX248244	BX248244 Human DNA
42 c	423	29.1	90732	9	AL670886	AL670886 Human DNA
43 c	423	29.1	100000	9	AP000503S2	AP000503S2 Homo sapi
44	423	29.1	136493	9	AL662899	AL662899 Human DNA
45	423	29.1	163682	9	HSDJ71I17	AL133713 Human DNA

ALIGNMENTS

RESULT 1

```
BD251986
LOCUS       BD251986                858 bp    DNA    linear   PAT 17-JUL-2003
DEFINITION  Screen method.
ACCESSION   BD251986
VERSION     BD251986.1  GI:33061756
KEYWORDS    JP 2002535001-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 858)
  AUTHORS   Vallance,P.J.T., Leiper,J.M., Whitley,G.S.J. and Charles,I.G.
  TITLE     Screen method
  JOURNAL   Patent: JP 2002535001-A 1 22-OCT-2002;
            UNIVERSITY COLLEGE LONDON
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002535001-A/1
            PD   22-OCT-2002
            PF   26-JAN-2000 JP 2000596131
            PR   26-JAN-1999 GB   9901705.5, 04-JUN-1999 GB   9913066.8 PI
            PATRICK JOHN THOMPSON VALLANCE,JAMES MITCHELL LEIPER,GUY ST PI
            JOHN WHITLEY,
            PI   IAN GEORGE CHARLES
            PC   C12N15/09,A01K67/027,A61K31/199,A61K38/00,A61K45/00,A61P3/06,
            PC   A61P9/02,
            PC   A61P9/04,A61P9/10,A61P9/12,A61P13/12,A61P25/06,A61P25/18,  PC
            A61P25/28,
            PC   A61P29/00,A61P31/04,A61P31/04,A61P35/00,C07K16/40,C12N1/15,C12N1/19,  PC
            C12N1/21,
            PC   C12N5/10,C12N9/78,C12Q1/02,C12Q1/34,G01N33/15,G01N33/50 PC
            ,C12N15/00,C12N5/00,
            PC   A61K37/02
            CC   Screen method
            FH   Key             Location/Qualifiers
            FT   CDS             (1)..(858).
FEATURES             Location/Qualifiers
     source          1..858
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                  6.53e-124
Score:                      1454.00
Percent Similarity:         100.00%         Length:         858
Best Local Similarity:      100.00%         Matches:        285
Query Match:                100.00%         Conservative:     0
DB:                                         Mismatches:       0
                                            Indels:           0
                                            Gaps:             6
```

US-09-889-733B-2 (1-285) x BD251986 (1-858)

```
Qy    1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValArgAla    20
Db    1 ATGGCCGGCCTCGGCCACCCCTCGGCCTTCGGCCGGGCCACCCACGCCGTGCGGGCG     60

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaGlyLyGluValAspVal     40
Db   61 CTACCCGAGTCGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC   120

Qy   41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValSerLysGlyLeuGlyLeu     60
Db  121 GCCCGGGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGTCGAAGGCAGCCAAGCTGGGGCTG   180

Qy   61 GlnValValGluGluThrLeuProAlaAspGluSerLeuProAspCysProValAspVal   80
Db  181 CAGGTGGTTGGAGGAACGGCAGCCTCCGGACTGCCTCTTCGTGAGGACGTG   240

Qy   81 AlaAlaValCysGluGluThrAlaLeuIleThrArgProValAlaProSerArgArgLys   100
Db  241 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGGCCGGCCCGAGCCGAGGAAG   300

Qy  101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys   120
```

```
Db  301 GAGGTTGACATGATGAAGGAAGCCTTGGAGAAGCTCCAGCTCAATATAGTAGAGATGAAA   360

Qy  121 AspGluAsnAlaThrLeuAspThrGlyGlyAspValLeuPheThrLyArgGluPheVal   140
Db  361 GATGAAAATGCCACTTTAGATGCGGGAGATGTTTTATTCACAGGCAGAGAATTTTGTG   420

Qy  141 GlyLeuSerLyArgThrAsnGlnArgGlyLyAlaGluIleLeuLeuAlaAspThrPheLyAsp   160
Db  421 GGCCTTCCAAAGGACAAAATCACGAGTGCTGAAAATCTTGCTGATACTTTTAAGGAC   480

Qy  161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLySerPheCysSerMet   180
Db  481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGTTGCATTGAAGAGTTTCTCGCAGCATG   540

Qy  181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLyAlaLeuLyIle   200
Db  541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC   600

Qy  201 MetGlnMetSerAspHisArgTyrAspLysLeuThrValProGlnAspAspIleAlaAla   220
Db  601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA   660

Qy  221 AsnCysIleTyrLeuAsnIleProAsnLyGlyHisValLeuLeuHisArgThrProGlu   240
Db  661 AACTGTATATATCTAAATATCCCAACAAAGGGCCACGTCTTGCTGCACCGAACCCCGAA   720

Qy  241 GluTyrProGluSerAlaLyValTyrGluLySerLeuLysAspHisMetLeuIleProVal   260
Db  721 GAGTATCCAGAAAGTGCAAAGTTTATGAGAACTGAAGGACCATATGCTGATCCCCGTG   780

Qy  261 SerMetSerGluLeuGluLyValAspGlyLeuLeuThrCysCysSerValLeuIleAsn   280
Db  781 AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC   840

Qy  281 LyLyValAspSer   285
Db  841 AAGAAGTAGACTCC   855
```

```
RESULT 2
AX032835
LOCUS       AX032835                858 bp    DNA    linear   PAT 21-SEP-2000
DEFINITION  Sequence 1 from Patent WO0044888.
ACCESSION   AX032835
VERSION     AX032835.1  GI:10279809
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Leiper,J.M., Whitley,G.S., Charles,I.G. and Vallance,P.J.
  TITLE     screen method
  JOURNAL   Patent: WO 044888-A 1 03-AUG-2000;
            LEIPER JAMES MITCHELL (GB) ; WHITLEY GUY ST JOHN (GB) ; UNIV LONDON
            (GB) ; CHARLES IAN GEORGE (GB) ; VALLANCE PATRICK JOHN THOMPSON
            (GB)
FEATURES             Location/Qualifiers
     source          1..858
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1..858
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09947.1"
                     /db_xref="GI:10279810"
                     /db_xref="REMTREMBL:CAC09947"
                     /translation="MAGLGHPSAFGRATHAVVRALIPESLCQHALRSAKGEVDVARAE
                     RQHQLYVGVLGSKLGLQVVEELPADESLPDCVVEDVAVVCEETALITRPGAPSRRKEV
                     DMMKEALEKLQLNIVEMKDENATLDGGDVLFTGREFVGLSKRTNQRGAEILADTFKD
                     YAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALKIMQMSDHRYDKLTVPDDI
                     AANCIYLNIPNKGHVLLHRTPEEYPESAKVYEKLKDHMLIPVSMSELEKVDGLLTCCS
```

```
RESULT 3
LOCUS       BC033680                1624 bp    mRNA    linear   PRI 07-OCT-2003
DEFINITION  Homo sapiens dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
            clone MGC:45161 IMAGE:5189970), complete cds.
ACCESSION   BC033680
VERSION     BC033680.1  GI:21707414
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1624)
AUTHORS     Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE       Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE     22388257
PUBMED      12477932
REFERENCE   2  (bases 1 to 1624)
AUTHORS     Strausberg,R.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 68 Row: m Column: 1
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 6912327.
FEATURES             Location/Qualifiers
     source          1..1624
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

```
ORIGIN          "VLINKKVDS"

Alignment Scores:
Pred. No.:
Score:                   6.53e-124     Length:            858
Percent Similarity:      1454.00       Matches:           285
Best Local Similarity:   100.00%       Conservative:        0
Query Match:             100.00%       Mismatches:          0
DB:                      100.00%       Indels:              0
                                       Gaps:                6

US-09-889-733B-2 (1-285) x AX032835 (1-858)

Qy     1  MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla   20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  ATGGCCGGCCTCGGCCACCCCTCCGCCTTCGGCCGGGCCACCCACGCCGTGGTGCGGGCG   60

Qy    21  LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValValAspVal   40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CTACCCGAGTCGCTCTGCCAGCACGCGCTGAGGAGCGCCAAGGGCGAGGAGGTGGACGTC  120

Qy    41  AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLeuLysGlyLeu   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGTGGGGCTG  180

Qy    61  GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal   80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CAGGTGGTGGAGCTGCCCGGCCGAGAGCCTTCCCGACTGCGTCTTCGTGGAGGACGTG  240

Qy    81  AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys  100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GCCGTGGTGTGCGAGGAGACGGCGCCCATCACCCGACCCGGGGCGCCGAGCCGGAGGAAG  300

Qy   101  GluValAspMetMetLysMetGluAlaLeuGlyLysGlnLeuAsnIleValGluMetLys  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GAGGTTGACATGATGAAGATGGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA  360

Qy   121  AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal  140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GATGAAAATGCAACTTTAGATGCCGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTG  420

Qy   141  GlyLeuSerLysArgThrArgGlnArgGlyAlaGlyAlaGluIleLeuAlaAspThrPheLysAsp  160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GGCCTTTCCAAAAGGACAAATCACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC  480

Qy   161  TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG  540

Qy   181  AlaGlyProAsnLeuIleAlaIleGlySerSerGlnSerAlaGlnLysAlaLeuLysIle  200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC  600

Qy   201  MetGlnGlnMetSerAspHisArgTyrAspLysThrValProAspAspAspIleAlaAla  220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  ATGCAACAGATGAGTGACCACCGCTACGACAAAACTGTCCCTGATGACATAGCAGCA  660

Qy   221  AsnCysIleTyrLeuAsnIleProAsnLysValGlyHisValLeuLeuHisArgThrProGlu  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  AACTGTATATATCTAAATATCCCAACAAAGGCACGTCTTGCTGCACCGGAACCCCGGAA  720

Qy   241  GluTyrProGluSerArgAlaLysValTyrGluLysTyrLeuLysAspHisMetLeuIleProVal  260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG  780

Qy   261  SerMetSerGluGluGlyLysValValAspGlyLeuLeuThrCysSerValLeuIleAsn  280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  AGCATGTCTGAACTGGAAGGTGGATGGGCTTCCACCTGCTCCAGTTTTAATTAAC  840

Qy   281  LysLysValAspSer  285
          |||||||||||||||
Db   841  AAGAAGGTAGACTCC  855
```

```
Db    520  GGCCTTCCAAAAGGACAAATCACGAGGTGCTGAAATCTTCGCTGATACTTTAAGGAC  579
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    161  TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet  180
Db    580  TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTCTGCAGCATG  639
Qy    181  AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle  200
Db    640  GCTGGGCCTAACCTGATCGCAATTGGGCTTAGTAGATCTGCACAGAGAGGCCCTTAAGATC  699
Qy    201  MetGlnMetSerAspHisArgTyrAspLysLeuThrValProAspArgAspAspIleAlaAla  220
Db    700  ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA  759
Qy    221  AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuLeuHisArgThrProGlu  240
Db    760  AACTGTATATATCTAAATATCCCAACAAAGGGCACGTTCTTGCTGCACCAGAACCCCGAA  819
Qy    241  GluTyrProGluSerAlaLysValTyrGluLysLeuGlyLysAspHisMetLeuIleProVal  260
Db    820  GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG  879
Qy    261  SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn  280
Db    880  AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTAATTAAC  939
Qy    281  LysLysValAspSer  285
           |||||||||||||||
Db    940  AAGAAAGTAGACTCC  954
```

```
RESULT 4
BD094073
LOCUS       BD094073              1633 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Shear stress-responsive DNAs.
ACCESSION   BD094073
VERSION     BD094073.1  GI:22639661
KEYWORDS    WO 0125427-A/34.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1633)
AUTHORS     Nojima,H., Yoshisue,H., Obayashi,M., Ota,T., Kawabata,A.,
            Sakurada,K., Kuga,T., Sekine,S., Nakamura,Y. and Sugai,S.
TITLE       Shear stress-responsive DNAs
JOURNAL     Patent: WO 0125427-A 34 12-APR-2001;
            KYOWA HAKKO KOGYO CO LTD,HIROSHI NOJIMA,HAJIME YOSHISUE, MASAYA
            OBAYASHI, TOSHIO OTA,AYAKO KAWABATA,KAZUHIRO SAKURADA,TETSURO KUGA,
            SUSUMU SEKINE, YUSUKE NAKAMURA, SUMIO SUGANO
COMMENT     OS   Homo sapiens (human)
            PN   WO 0125427-A/34
            PD   12-APR-2001
            PF   02-OCT-2000 WO 2000JP006840
            PR   01-OCT-1999 JP 99P 280976
            PI   HIROSHI NOJIMA,HAJIME YOSHISUE,MASAYA OBAYASHI,TOSHIO OTA, PI
            PI   AYAKO KAWABATA,
            PI   KAZUHIRO SAKURADA,TETSURO KUGA,SUSUMU SEKINE,YUSUKE NAKAMURA,
            PI   SUMIO SUGANO
            PC   C12N15/12,C07K14/435,C07K16/18,C12P21/02,C12Q1/68,A61K38/00,
            PC   A61K39/395,
            PC   A61K48/00,A61P9/10,G01N33/50,G01N33/53
            PC
            FH   Key             Location/Qualifiers
            FT                   (323)...(1177).
FEATURES             Location/Qualifiers
     source          1..1633
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
                 /clone="MGC:45161 IMAGE:5189970"
                 /tissue_type="Colon, Kidney, Stomach, adult, whole pooled"
                 /clone_lib="NIH_MGC_116"
                 /lab_host="DH10B"
                 /note="Vector: pCMV-SPORT6"
     gene        1..1624
                 /gene="DDAH1"
                 /note="synonym: DDAH"
                 /db_xref="LocusID:23576"
                 /db_xref="MIM:604743"
     CDS         100..957
                 /codon_start=1
                 /product="DDAH1 protein"
                 /protein_id="AAH33680.1"
                 /db_xref="GI:21707415"
                 /db_xref="LocusID:23576"
                 /translation="MAGLGHFAAFGRATHAVVRALPESLGQHALRSAKGEVDVARAE
                 RQHQLYVGLGSKLGLQVVELPADESLPDCVVPEDVAVVCEETALITRPGAPSRRKEV
                 DMMKEALEKLQLNIVEMKDENATLDGGDVLFTGREFPVGLSKRTNQRGAEILADTFKD
                 YAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDI
                 AANCIYLNIPNKGHVLLHRTPEYPESAKVYEKLKDHMLIPVSMSELEKVDGLLTCCS
                 VLINKKVDS"
     misc_feature 133..936
                 /note="Amidinotransf; Region: Amidinotransferase. This
                 family contains glycine (EC:2.1.4.1) and inosamine
                 (EC:2.1.4.2) amidinotransferases, enzymes involved in
                 creatine and streptomycin biosynthesis respectively. This
                 family also includes arginine deiminases, EC:3.5.6.
                 These enzymes catalyse the reaction: arginine + H2O <=>
                 citrulline + NH3. Also found in this family is the
                 Streptococcus anti tumor glycoprotein"
                 /db_xref="CDD:pfam02274"
```

```
ORIGIN

Alignment Scores:
Pred. No.:              3.62e-122       Length:            1624
Score:                  1439.00         Matches:           283
Percent Similarity:     99.65%          Conservative:      1
Best Local Similarity:  99.30%          Mismatches:        1
Query Match:            98.97%          Indels:            0
DB:                     9               Gaps:              0

US-09-889-733B-2 (1-285)  x  BC033680 (1-1624)

Qy     1  MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla  20
          ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db   100  ATGGCCGGGCTCGGCCACCCGTCGGCCTTTCGGCCGGCCACCCACGCCGTGGTGCGGGCG  159

Qy    21  LeuProGluSerLeuCysGlnHisAlaLeuArgLeuAlaLysGlyGluValAspVal  40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   160  CTACCCGAGTCGCTGTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC  219

Qy    41  AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   220  GCCCGCGCCGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG  279

Qy    61  GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal  80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   280  CAGGTGGTCGAGCTGCCCGGCCGACGAGAGCCCTTCCGGACTGCGTCTTCGTGGAGGACGTG  339

Qy    81  AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys  100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   340  GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCCCCAGCCCGGAGGAAG  399

Qy   101  GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   400  GAGGTTGACATGATGAAAGAAGCATTAGAAAACTTCAGCTCAATAATAGTAGAGATGAAA  459

Qy   121  AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal  140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   460  GATGAAAATGCAACTTTAGATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTG  519

Qy   141  GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp  160
```

```
Pred. No.:              3.64e-122      Length:          1633
Score:                  1439.00        Matches:         283
Percent Similarity:     99.65%         Conservative:    1
Best Local Similarity:  99.30%         Mismatches:      1
Query Match:            98.97%         Indels:          0
DB:                                    Gaps:            0
```

US-09-889-733B-2 (1-285) x BD094073 (1-1633)

```
Qy    1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
          |||||||||||||||||||||||||::::|||||||||||||||||||||||
Db  323 ATGGCCGGGCTCGGACATCCTTCGGCCCCCGGCCTTCGGCCGTGGTGCGGGCG 382

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  383 CTACCCGAGTCGCTGCTCGGCCAGCACGCGCTGAGAAGCGCCAAGGGCCAGGAGGTGGACGTC 442

Qy   41 AlaArgAlaAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlyGlyLysLeuGlyLeu 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  443 GCCCGCGCCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCCAGCAAGCTGGGGCTG 502

Qy   61 GlnValValGluLeuProAlaAspGluSerLeuProAlaAspCysValPheValGluAspVal 80
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  503 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTCTTCCCGACTGCGGAGGACGTG 562

Qy   81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  563 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGG 622

Qy  101 GluValAspMetMetLysGlyValLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  623 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 682

Qy  121 AspGluAsnAlaThrLeuAspGlyAspValLeuPheThrGlyArgGluPhePheVal 140
          |||||||||||||||||||||||||||||||||||||||||||||
Db  683 GATGAAAATGCAACTTTAGATGGCGAGATGTTTTATTCACAGGCAGAGAATTTTTTTGTG 742

Qy  141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
          |||||||||||||||||||||||||||||||||||||||||||||
Db  743 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 802

Qy  161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
          |||||||||||||||||||||||||||||||||||||||||||||
Db  803 TATGCAGTTCCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 862

Qy  181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
          |||||||||||||||||||||||||||||||||||||||||||
Db  863 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 922

Qy  201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspAspIleAlaAla 220
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  923 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 982

Qy  221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  983 AACTGTATATATTCTAAATATCCCCAACAAAGGCCACGTTCTTGCTGCACCGAACCCCGAA 1042

Qy  241 GluTyrProGluSerAlaLysValTyrGluLysValAspGlyLeuLeuThrCysCysSer 260
          |||||||||||||||||||||||||||||||||||||||||||||
Db 1043 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGGACGGACCATATGTGCTGATCCCCGTG 1102

Qy  261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerSerMetSerLeuLeuAsn 280
          ||||||||||||||||||||||||||||||||||||||||||||
Db 1103 AGCATGTCTGAACTGCAAAAGGTTGGATGGGGTCTACTGGCCTGCTGAGTTTTAATTAAC 1162

Qy  281 LysLysValAspSer 285
          ||||||||||||
Db 1163 AAGAAAGTAGACTCC 1177
```

```
ACCESSION   AB001915
VERSION     AB001915.1  GI:4160665
KEYWORDS    NG,NG-dimethylarginine dimethylaminohydrolase.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
AUTHORS     Kimoto,M., Miyatake,S., Sasagawa,T., Yamashita,H., Okita,M.,
            Oka,T., Ogawa,T. and Tsuji,H.
TITLE       Purification, cDNA cloning and expression of human
            NG,NG-dimethylarginine dimethylaminohydrolase
JOURNAL     Eur. J. Biochem. 258 (2), 863-868 (1998)
MEDLINE     99089653
PUBMED      9874257
REFERENCE   2  (bases 1 to 1633)
AUTHORS     Kimoto,M.
TITLE       Direct Submission
JOURNAL     Submitted (15-MAR-1997) Masumi Kimoto, Okayama Prefectural
            University, Faculty of Health and Welfare Science, Department of
            Nutritional Science; Kuboki 111, Soja, Okayama 719-1197, Japan
            (E-mail:kimoto@fhw.oka.-pu.ac.jp, Tel:0866-94-2158,
            Fax:0866-94-2158)
FEATURES             Location/Qualifiers
     source          1..1633
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="Kidney"
     CDS             323..1180
                     /EC_number=3.5.3.18"
                     /codon_start=1
                     /product="NG,NG-dimethylarginine dimethylaminohydrolase"
                     /protein_id="BAA37117.1"
                     /db_xref="GI:4160666"
                     /translation="MAGLGHPAAFGRATHAVVRALPESLGQHALRSAKGEEVDVARAE
                     RQHQLYVGVLGSKLGLQVVELPADESLPDCVFVEDVAVVCEETALITRPGAPSRRKEV
                     DMMKEALEKLQLNIVEMKDENATLDGDVLFTGREFFVGLSKRTNQRGAEILADTFKD
                     YAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDI
                     AANCIYLNIPNKGHVLLHRTPEEYPESAKVYEKLKDHMLIPVSMSELEKVDGLLTCCS
                     VLINKKVDS"

ORIGIN
```

```
Alignment Scores:
Pred. No.:              3.64e-122      Length:          1633
Score:                  1439.00        Matches:         283
Percent Similarity:     99.65%         Conservative:    1
Best Local Similarity:  99.30%         Mismatches:      1
Query Match:            98.97%         Indels:          0
DB:                                    Gaps:            0
```

US-09-889-733B-2 (1-285) x AB001915 (1-1633)

```
Qy    1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
          |||||||||||||||||||||||::::||||||||||||||||||||||||
Db  323 ATGGCCGGGCTCGGACATCCTTCGGCCCCCGGCCTTCGGCCGTGGTGCGGGCG 382

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  383 CTACCCGAGTCGCTGCTCGGCCAGCACGCGCTGAGAAGCGCCAAGGGCCAGGAGGTGGACGTC 442

Qy   41 AlaArgAlaAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  443 GCCCGCGCCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCCAGCAAGCTGGGGCTG 502

Qy   61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  503 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTCTTCCCGACTGCGGAGGACGTG 562

Qy   81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  563 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCCGAGCCGGAGGAAG 622
```

```
RESULT 5
AB001915
LOCUS       AB001915                1633 bp    mRNA    linear   PRI 21-SEP-2000
DEFINITION  Homo sapiens mRNA for NG,NG-dimethylarginine
```

```
Query Match:           98.97%      Indels:        0
DB:                    6           Gaps:          0

US-09-889-733B-2 (1-285) x AX548046 (1-4208)

Qy      1  MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla  20
Db    368  ATGGCCGGGCTCGGCCACCCCTCGGCCGGCCGGGCCACCCACGCCGCCGGTCGCGGGCG  427

Qy     21  LeuProGluSerLeuCysGlnHisAlaAlaLeuArgSerAlaLysGlyGluValAspVal  40
Db    428  CTACCCGAGTCGCTCGGCCACGCCGCGTGAGAAGCGGCAAGGGCGAGGAGGTGGACGTC  487

Qy     41  AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLysGlyLeu  60
Db    488  GCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG  547

Qy     61  GlnValValGluLeuProAlaAspCysValPheGluGluValAspVal  80
Db    548  CAGGTGGTGGAGCTGCCGGCAGAGGCCTTCCGGACACTGGCGTCTCGTGGAGGACGTG  607

Qy     81  AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys  100
Db    608  GCCGTGGTGTGCGGAGGACGACGGCCCTCATCACCCGGCCGGGCGCCCCGAGCCGGAAG  667

Qy    101  GlyValValAspMetMetLysGluAlaLeuGlnLeuAsnIleValGluMetLys  120
Db    668  GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGTCACATATAGTAGATGAA  727

Qy    121  AspGluAsnAlaThrLeuAspGlyGlyAspAspArgGlyArgGluPhePheVal  140
Db    728  GATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG  787

Qy    141  GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAsp  160
Db    788  GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC  847

Qy    161  TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet  180
Db    848  TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG  907

Qy    181  AlaGlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIle  200
Db    908  GCTGGCCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC  967

Qy    201  MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla  220
Db    968  ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA  1027

Qy    221  AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu  240
Db   1028  AACTGTATATATCTAAATATCCCAACAAAGGGCCACGTCTTGCTGCACCGGACCCCGGAA  1087

Qy    241  GluTyrProGluSerAlaLysValTyrGluLysLysAspHisMetLeuProVal  260
Db   1088  GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAGAAAGGACCATATGCTGATCCCCGTG  1147

Qy    261  SerMetSerGluLeuGluLysCysSerSerValLeuIleAsn  280
Db   1148  AGCATGTCTGAACTGGAAAAGTGCAGTGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC  1207

Qy    281  LysLysValAspSer  285
Db   1208  AAGAAGTAGACTCC  1222
```

RESULT 7
BC034505
LOCUS   BC034505   1120 bp   mRNA   linear   ROD 25-NOV-2003
DEFINITION   Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
               clone MGC:30267 IMAGE:5134831), complete cds.
ACCESSION   BC034505
VERSION   BC034505.1 GI:21961623
KEYWORDS   MGC.

```
Qy    101  GluValAspMetMetLysGluAlaLeuGlnLeuAsnIleValGluMetLys  120
Db    623  GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGTCACATATAGTAGATGATGAA  682

Qy    121  AspGluAsnAlaThrLeuAspGlyGlyAspAspArgGlyArgGluPhePheVal  140
Db    683  GATGAAAATGCAACTTTAGATGGGCGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG  742

Qy    141  GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAsp  160
Db    743  GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC  802

Qy    161  TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet  180
Db    803  TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG  862

Qy    181  AlaGlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIle  200
Db    863  GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC  922

Qy    201  MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla  220
Db    923  ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA  982

Qy    221  AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu  240
Db    983  AACTGTATATATCTAAATATCCCCACAAAGGGCCACGTCTTGCTGCACCGGACACCCGGAA  1042

Qy    241  GluTyrProGluSerAlaLysValTyrGluLysLysAspHisMetLeuProVal  260
Db   1043  GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAGAAACTGAAGGACCCATATGCTGATCCCCGTG  1102

Qy    261  SerMetSerGluLeuGluLysCysSerSerValLeuIleAsn  280
Db   1103  AGCATGTCTGAACTGGAAAAGTGCAGTGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC  1162

Qy    281  LysLysValAspSer  285
Db   1163  AAGAAGTAGACTCC  1177
```

RESULT 6
AX548046
LOCUS   AX548046   4208 bp   DNA   linear   PAT 26-NOV-2002
DEFINITION   Sequence 22 from Patent WO02066654.
ACCESSION   AX548046
VERSION   AX548046.1 GI:25813142
KEYWORDS
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Astromoff,A., Au-Young,J., Baughn,M.R., Ding,L., Duggan,B.M.,
               Forsythe,I.J., Gietzen,K.J., Griffin,J.A., Lee,E.A., Lu,Y.,
               Richardson,T.W., Ring,H.Z., Sanjanwala,M.M., Swarnakar,A.,
               Walia,N.K., Warren,B.A., Xu,Y., Yue,H. and Zebarjadian,Y.
  TITLE   Drug metabolizing enzymes
  JOURNAL   Patent: WO 02066654-A 22 29-AUG-2002;
               Incyte Genomics, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..4208
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7671089CB1"
ORIGIN
Alignment Scores:
Pred. No.:   1..4208
Score: 1.29e-121   Length: 4208
Percent Similarity: 1439.00   Matches: 283
Best Local Similarity: 99.65%   Conservative: 1
                          99.30%   Mismatches: 1

```
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1120)
AUTHORS     Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE       Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE     22338257
PUBMED      12477932
REFERENCE   2 (bases 1 to 1120)
AUTHORS     Strausberg,R.
TITLE       Direct Submission
JOURNAL     Submitted (24-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 41 Row: O Column: 4
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis. Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1..1120
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:30267 IMAGE:5134831"
                     /tissue_type="liver, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1..1120
                     /gene="Ddah1"
                     /note="synonym: 2510015N06Rik"
                     /db_xref="LocusID:69919"
                     /db_xref="MGI:1916469"
     CDS             122..979
                     /codon_start=1
                     /product="dimethylarginine dimethylaminohydrolase 1"
                     /protein_id="AAH34505.1"
                     /db_xref="GI:21961624"
                     /db_xref="LocusID:69919"
                     /translation="MAGLGHFSAFGRATHAVVRAPPESLCRHALRSQGEVDFPARAE
                     RQHELYVGVLGSKLGLGQVVQLPADESLPDCVVRDVAVVCBETALITRPGAPSRRKEV
                     DMMKEALEKLQLNIVEMKDENATLDGDVLFTGREFFVGLSKRTNQRGAEILADTFKD
                     YAVSTVPVADSLHLKSPCSMAGPNLIAIGSSRSAQKALKIMQCMSDHRYDKLTVPDDM
                     AANCIYLNIPSKGHVLLHRTPEYPESAKVYEKLKDHLLIFVSNSEMEKVDGLLTCCS
                     VFINKKIDS"
     misc_feature    155..958
                     /note="Amidinotransf; Region: Amidinotransferase. This
                     family contains glycine (EC:2.1.4.1) and inosamine
                     (EC:2.1.4.2) amidinotransferases, enzymes involved in
                     creatine and streptomycin biosynthesis respectively. This
                     family also includes arginine deiminases, EC:3.5.3.6.
                     These enzymes catalyse the reaction: arginine + H2O <=>
                     citrulline + NH3. Also found in this family is the
                     Streptococcus anti tumour glycoprotein"
                     /db_xref="CDD:pfam02274"

ORIGIN

Alignment Scores:
Pred. No.:                 1                    Length:               1120
Score:                     8.32e-118            Matches:              1389.00
Percent Similarity:        97.89%               Conservative:         269
Best Local Similarity:     94.39%               Mismatches:           10
Query Match:               95.53%               Indels:               6
DB:                                             Gaps:                 0

US-09-889-733b-2 (1-285) x BC034505 (1-1120)

Qy    1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 ATGGCCGGCCTCGGCCACCCCTCCGCCTTCGGCCGGGCCACCCACGCCGTGGTGCGGGCT 181

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAspVal 40
           ||||||||||||| :::|||| |||||||||| |||||||||||| :::||||||||
Db  182 CCGCCCGAGTCCCTGTGCCGCCACGCGCTGAGGCGCCTCCAGGGGCGAGAGGTGGATTTC 241

Qy   41 AlaArgAlaGluArgGlnHisGlnHisLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
           |||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db  242 GCTCGCCGCCAGCGCCAGCACGAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCCTG 301

Qy   61 GlnValGluLeuProAlaAspGluSerLeuProAspCysValValArgAspVal 80
           |||||||:||:::| ||||||||||||||||||||||||||||||||||||||||
Db  302 CAGGTGGTGCAGCTGCCCGCGGACGAGAGCCTGCCCGACTGCGTGTTCGTGGAGGACGTG 361

Qy   81 AlaValValCysGlyGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
           ||||||||:::||||||| |||||||:::||||||||||||||||||||||||||||||
Db  362 GCCGTCGTGTGCGGAGAGACGGCGCTGATCACCAGGCCCGGGGCGCCCAGCCGCCAGGAAG 421

Qy  101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  422 GAGGTTGACATGATGAAGAGAGGCTTTGGAAAACTTCAGCTCAACATAGTAGAGATGAAA 481

Qy  121 AspGluAsnAlaThrLeuAspGlyGlyValLeuPheThrGlyArgGluPheThrValFheVal 140
           ||:|||||:::||||||||||||| |||||| |||||||||||||||||||||||||TTG
Db  482 GATGAAAATGCAACTTTGGATGGTGGGACGTCCTATTCACAGGCAGGAGAATTTTTTGTG 541

Qy  141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  542 GGCCTTTCCAAAGAACAAAATCAACCGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 601

Qy  161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  602 TACGCAGTCTCTACAGTCCCCTGTGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCATG 661

Qy  181 AlaGlyProAsnLeuIleAlaIleGlySerSerArgSerAlaGlnLysAlaLeuLysIle 200
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db  662 GCCGGACCCAACCTGATTGCAATAGGGTCCAGCAGATCTGCACAGAAGGCCCTCAAGATC 721
```

Qy 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
        ||||||||||||||||||||||||||||||||||||||||||||||| |||:||
Db 722 ATGCAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGACATGGCCGCC 781

Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
        ||||| | ||||| || ||||:||||||||:||||| |||||||||||||||||
Db 782 AACTGCATATATCTAAATATCCCCAGCAAAGGGCATGTCTTGCTGCACCGAACCCCAGAA 841

Qy 241 GluTyrProGluSerAlaLysValTyrGluLysAspHisMetLeuIleProVal 260
        ||||||||||||||||||||||||||||:||||||||||||||||||||||
Db 842 GAGTACCCAGAAAGCGCAAAGGTCTATGAGAAACTCAAGGACCATCTACTGATCCCTGTG 901

Qy 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
        ||| ::|||||||||||||||||||||||||||||||||||||||||||||||||
Db 902 AGCAACTCGGAGATGAAAAGGTGGACGGGCTTGCTCACCTGCTGCTCCGTTTTTATTAAC 961

Qy 281 LysLysValAspSer 285
        |||:::||||||
Db 962 AAGAAGATAGACTCC 976

RESULT 8
AX683109
LOCUS   AX683109   3008 bp   DNA   linear   PAT 29-MAR-2003
DEFINITION Sequence 83 from Patent EP1279744.
ACCESSION   AX683109
VERSION   AX683109.1 GI:29370146
KEYWORDS
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Brooksbank,R.A., Dixon,A.K., Lee,K. and Pinnock,R.D.
  TITLE   Identification and use of molecules implicated in pain
  JOURNAL   Patent: EP 1279744-A 83 29-JAN-2003;
            WARNER-LAMBERT COMPANY (US)
FEATURES   Location/Qualifiers
     source   1..3008
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
                     /note="N-G, N-G dimethylarginine dimethylaminohydrolase"

ORIGIN

Alignment Scores:
Pred. No.:   2.56e-116   Length: 3008
Score: 1379.00   Matches: 268
Percent Similarity: 96.84%   Conservative: 8
Best Local Similarity: 94.04%   Mismatches: 9
Query Match: 94.84%   Indels: 0
DB: Gaps: 0

US-09-889-733B-2 (1-285) x AX683109 (1-3008)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 432 ATGGCCGGCGCTCAGCCACCCCTCGGTCTTGGCCGCGCCACCCACGCCGTGGTCCGGGCT 491

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuAArgSerAlaLysGlyGluGluValAspVal 40
        ||||| |||||||||:||::|||| |:|||:|||||||||||||||||||||
Db 492 CCGCCCAGTCCCTGCGCGCAGCGCCTGAGCGCTGGAGGCGAAGAGGTGGATTTC 551

Qy   41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyLeuTyrValGlySerLysLysGlyLeu 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 552 GCTCGGCGTGAGCGCCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCTGGGGCTG 611

Qy   61 GlnValValGluLeuProAlaAlaGluSerLeuProAspCysValPheValGluAspVal 80
        ||||||||:::|||||||||||||||||||||||||||||||||||||:||
Db 612 CAGGTGGTCGAGCTGCCCGGCCGCCGAGAGCCTGACCTGGTCCGTGTTCGTGGAGGACGTG 671

Qy   81 AlaValValCysGluGluThrAlaIleIleThrArgProGlyAlaLeuAlaProSerArgArgLys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||

```
                         Qy   201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
                                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                         Db  1032 ATGCAACAGATGAGTGACCACCGTTATGACAAGCTCACTGTACCGGACGACATGGCCGCC 1091

                         Qy   221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
                                  ||||||||||||||||||||  |||||||||||||||||||||||||||||||| |||||
                         Db  1092 AACTGTATATATTTAAATATCCCCAGCAAAGGGCATGTCTTGCTGCACCGAACCCCAGAA 1151

                         Qy   241 GluTyrProGluSerAlaLysValTyrGluLysLysAspHisMetLeuIleProVal 260
                                  ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
                         Db  1152 GAGTACCCAGAAAGCGCAAAGGTTTATGAGAAGAAGTCAAGGACCATCTACTGATCCCTGTG 1211

                         Qy   261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
                                  |||||||| ||||||| ::::|||||||||||||| ||||||||||||||| ::||||
                         Db  1212 AGCAATTCTGAGATGCAAAGGTGACGGCGGCTTGCTCACCTGCTGCTCCGTTTTTATTAAC 1271

                         Qy   281 LysLysValAspSer 285
                                  |||||| ||||||
                         Db  1272 AAGAAGACAGACTCT 1286
```

```
RESULT 10
BC043235
LOCUS       BC043235              3737 bp    mRNA    linear   PRI 30-SEP-2003
DEFINITION  Homo sapiens dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
            clone IMAGE:5295723), partial cds.
ACCESSION   BC043235
VERSION     BC043235.2  GI:34783628
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3737)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D. Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3737)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Sep 16, 2003 this sequence version replaced gi:28175756.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
```

```
770, Japan (E-mail:kimoto@nutr.med.tokushima-u.ac.jp,
    Tel:0886-33-7088, Fax:0886-33-7089)
FEATURES             Location/Qualifiers
     source          1..3008
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /sex="female"
                     /tissue_type="kidney"
     CDS             432..1289
                     /codon_start=1
                     /note="DDAH"
                     /product="N-G,N-G-dimethylarginine dimethylaminohydrolase"
                     /protein_id="BAA18993.1"
                     /db_xref="GI:1906800"
                     /translation="MAGLSHPSVFGRATHAVVRAPPESLCRHALRRSGBEVDFARAE
                     RQHQLYVGVLGSKLGLQVVQLPADESLPDCVFBDVAVVCEETALITRPGAPSRRKEV
                     DMMKEALEKLQLNIVEMKDENATLDGGDVLFTGREFFVGLSKRTNQRGAEILADTFKD
                     YAVSTVPVADSLHLKSFCSMAGPNLIAIGSSESAQKALKIMQCMSDHRYDKGTVPDDM
                     AANCIYLNIPSKGHVLLHRTPBEYPESAKVYEKLKDHLLIPVSNSEMEKVDGLLTCCS
                     VFINKKTDS"
```

```
ORIGIN
Alignment Scores:
Pred. No.:                                     3008
Score:            2.56e-116     Length:        3008
Percent Similarity:  1379.00    Matches:       268
Best Local Similarity:  96.84%  Conservative:  8
Query Match:      94.04%        Mismatches:    9
DB:               94.84%        Indels:        0
                  10            Gaps:          0
```

```
US-09-889-733B-2 (1-285) x D86041 (1-3008)

Qy     1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   432 ATGGCCGGCCTCAGCCACCCCTCGGTCTTCGGCCGGGCCACCCACGCCGTGGTGCGGGCT 491

Qy    21 LeuProGluSerLeuCysArgHisAlaLeuArgArgSerAlaGlyGluValAspVal 40
           ||||||||||||||||||||||  :::::||||||||||||||||||||||||||
Db   492 CCGCCCAGTCCCTGTGCCGGCACGGGCTGAGGCGCTCCCAGGGCAAGGAGGTGGATTTC 551

Qy    41 AlaArgAlaGluArgGlnHisHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   552 GCTCGCGCTGAGCGCCAGCACCAGCTCTACGTGGGGGTGCTGGGCAGCAAGCTGGGCGTG 611

Qy    61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   612 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTGCCTGACTGCGTGTTCGTGGAGGACGTG 671

Qy    81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   672 GCCGTGGTGTGCGAGGAGACCGCCCTCATCACCCGCCCGGGCGCCCCTAGCCGCCGGAAG 731

Qy   101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGlnMetLys 120
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   732 GAGGTTGCACATGCATGAAAGAGAGCCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAA 791

Qy   121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
           ||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   792 GATGAAAATGCAACCTTAGATGGTGGGGACGTCCATTCCACAGGCAGAGAGTTTTTGTG 851

Qy   141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   852 GGCCTTTCCAAAAGGACAAAATCAACCAGGTGCTGAGATCTTGGCTGATACTTTCAAGGAC 911

Qy   161 TyrAlaValSerThrValProValAlaAspGlyLeuHisHisLysLysSerPheCysSerMet 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   912 TACGCAGTTTCCACAGTCCCCGTGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCATG 971

Qy   181 AlaGlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIle 200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   972 GCTGGGCCCCAACCTGATCGCAATAGGGTCCAGTGAATCTGCGCAGAAGGCCCTCAAGATC 1031
```

```
Qy   80  ValAlaValValCysGluGluThrArgAlaLeuIleThrArgProGlyAlaProSerArgArg   99
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GTGGCCGTGGTGTGCGAGGAGACGAGCGGCCCTCATCACCCGAGCCGGGGGCCGAGCCGGAGG  180

Qy  100  LysGluValAspMetMetGlyAlaAlaGluGluLysLeuGlnLeuAsnIleValGluMet  119
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  AAGGAGGTTGACATGATGAAAGAAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATG  240

Qy  120  LysAspGluAsnAlaThrLeuAspGlyGlyAspValPheThrGlyArgGluPhePhe  139
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  AAAGATGAAAATGCAACTTAGATGGCGGAGATGTTTTATTCACAGGCCAGAGAATTTTTT  300

Qy  140  ValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLys  159
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  GTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGCAAATCTTGGCTGATACTTTTAAG  360

Qy  160  AspTyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSer  179
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GACTATGCAGTCTCCACAGTGCCAGTGGCCGACATGGGTTGCATTTGAAGAGTTTCTGCAGC  420

Qy  180  MetAlaGlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLys  199
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  ATGGCTGGGCCTAACCTGATCGCAATTGGCTCAGTGAATCTGCACAGAAGGCCCTTAAG  480

Qy  200  IleMetGlnGlnMetSerArgThrValProLysSerLeuThrValProAspAspIleAla  219
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  ATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCA  540

Qy  220  AlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro  239
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  GCAAACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCG  600

Qy  240  GluGluTyrProGluSerAlaLysValTyrGluLysAspHisMetIlePro  259
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  GAAGAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCC  660

Qy  260  ValSerMetSerGluLeuGluLysValAspHisGlyLeuLeuThrCysCysSerValLeuIle  279
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  GTGAGCATGTCTGAACTGGAAAAGGTGGATGGGGGTCTGCTGCTGTCAGTTTTAATT  720

Qy  280  AsnLysLysValAspSer  285
         ||||||||||||||||||
Db  721  AACAAGAAAGTAGACTCC  738
```

RESULT 11
AB095027
LOCUS   AB095027   1150 bp   mRNA   linear   VRT 27-JUN-2003
DEFINITION Gallus gallus mRNA for dimethylarginine dimethylaminohydrolase I,
            complete cds.
ACCESSION   AB095027
VERSION   AB095027.2 GI:32306458
KEYWORDS
SOURCE   Gallus gallus (chicken)
  ORGANISM Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1
  AUTHORS   Mishima,T., Hamada,T., Ui-Tei,K., Takahashi,F., Miyata,Y.,
            Imaki,J., Suzuki,H. and Yamashita,K.
  TITLE   The expression of DDAH I in the chick and rat embryo
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1150)
  AUTHORS   Mishima,T., Hamada,T., Ui-Tei,K., Takahashi,F., Miyata,Y.,
            Imaki,J., Suzuki,H. and Yamashita,K.
  TITLE   Direct Submission
  JOURNAL   Submitted (30-OCT-2002) Takuya Mishima, Nippon Medical School,
            Department of Anatomy; Sendagi 1-1-5, Bunkyo, Tokyo 113-8602, Japan
            (E-mail:mishima@nms.ac.jp, Tel:81-3-3822-2131(ex.5315))
COMMENT   On Jun 26, 2003 this sequence version replaced gi:30268647.
FEATURES   Location/Qualifiers
     source   1..1150
                     /organism="Gallus gallus"

BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 76 Row: j Column: 15.

```
FEATURES             Location/Qualifiers
     source          1..3737
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5295723"
                     /tissue_type="Testis"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
     gene            <1..3737
                     /gene="DDAH1"
                     /note="synonym: DDAH"
                     /db_xref="LocusID:23576"
                     /db_xref="MIM:604743"
     CDS             <1..741
                     /gene="DDAH1"
                     /codon_start=1
                     /product="DDAH1 protein"
                     /protein_id="AAH43235.2"
                     /db_xref="GI:34783629"
                     /db_xref="LocusID:23576"
                     /db_xref="MIM:604743"
                     /translation="VARAERQHQLYVGVLGSKLGLQVVELPADESLPDCVFVEDVAVV
                     CEETALITRPGAPSRRKRVDMMKEALEKLQLNIVEMKDENATLDGGDVLFTGREFFVG
                     LSKRTNQRGAEILADTFKDYAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALK
                     IMQQMSDHRYDKLTVPDDIAANCIYLNIPNKGHVLLHRTPEEYPESAKVYEKLKDHML
                     IPVMSELEKVDGLLTCCSVLINKKVDS"
     misc_feature    1..720
                     /gene="DDAH1"
                     /note="Amidinotransf; Region: Amidinotransferase. This
                     family contains glycine (EC:2.1.4.1) and inosamine
                     (EC:2.1.4.2) amidinotransferases, enzymes involved in
                     creatine and streptomycin biosynthesis respectively. This
                     family also includes arginine deiminases, EC:3.5.3.6.
                     These enzymes catalyse the reaction: arginine + H2O <=>
                     citrulline + NH3. Also found in this family is the
                     Streptococcus anti tumor glycoprotein"
                     /db_xref="CDD:pfam02274"
ORIGIN
```

Alignment Scores:
Pred. No.:
Score: 1.17e-104   Length: 3737
                        1253.00   Matches: 246
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 86.18%   Indels: 0
DB: 9   Gaps: 0

US-09-889-733B-2 (1-285) x BC043235 (1-3737)

```
Qy   40  ValAlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGly  59
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  GTGGCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGG  60

Qy   60  LeuGlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAsp  79
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  CTGCAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGAC  120
```

```
                    /mol_type="mRNA"
                    /db_xref="taxon:9031"
                    /tissue_type="embryos yolk sac"
                    /clone_lib="Chick Yolk Sac Lambda cDNA Library (Stratagene
                    #937406)"
                    /dev_stage="five-days-old chick"
    CDS             40..903
                    /codon_start=1
                    /product="dimethylarginine dimethylaminohydrolase I"
                    /protein_id="BAC76002.1"
                    /db_xref="GI:30268648"
                    /translation="MAGLGGGPAAFGRCTHAVVRALPESLCRQALRSTAGPEVDFARA
                    EREHQLYVGVLRGKLGLQVLELPADESLPDCVFVEDAAVVCEETALLTRPGAPSRRKE
                    VEAMKRVLESLNLNMVEMVDENATLDGGDVLFTGREFFFVGLSKRTNQRGAEILADTF
                    KDYAVSTVPVHDSLHLKSFCSMAGPNLIAIGSSEAAQKALKTMQQMSDHRYDKLITVPD
                    DAAANCIYLNIPSKGHVLLHRAPEYPESAKVFEKLKDHMLIPIANTELEKVDGALTC
                    CSVLINKTSEL"

ORIGIN
Alignment Scores:
Pred. No.:              2e-104
Score:                  1243.00
Percent Similarity:     91.17%            Length:         1150
Best Local Similarity:  86.57%            Matches:        245
Query Match:            85.49%            Conservative:   13
DB:                                       Mismatches:     13
US-09-889-733B-2 (1-285) x AB095027 (1-1150)    Indels:         2
                                          Gaps:           2
```

```
Qy    1 MetAlaGlyLeu---GlyHisProSerAlaPheGlyArgArgAlaThrHisAlaValValArg  19
          ||| ||| |||     ||||||||||||||  ||||||||| ||||| ||||| |||||||
Db   40 ATGGCCGGCCTGGGCGGCGGCCCCGCCTTCGGCCGCCGTACGCACGCCGTGGTGCGG       99

Qy   20 AlaLeuProGluSerLeuCysArgGlnHisAlaLeuArgSerAlaLysGlyGluValAsp     39
          |||||||||||||| || |||  |||  |:::  ||||||| |:|| |||
Db  100 GCGCTGCCCGAGTCGCTGTGCCGCCAGGCGTTGCGGAGCACGGCGGGCCCCGAGGTGGAC   159

Qy   40 ValAlaAlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGly  59
          ||||| || |  |||  ||| ||::::||||||||||||||||||||||||
Db  160 TTCGCTCCGCCGGAGCGGGAGCCATCAGCTGTACGTGGGCGTGCTACGGGGCAAGCTGGGG   219

Qy   60 LeuGlnValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAsp       79
          |:::|||| ||| ||||||||||||||  |||||||||||| |||||||||||||
Db  220 CTGCAGGTGGCTGGAGCTGCCGGCTGACGAGAGCCTCCCGACTGCGTCTTCGTGGAGGAC   279

Qy   80 ValAlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArg     99
          |||||| || |||||||||||||| |||  :||||||  |||  ||  ||
Db  280 GCGGCGGTAGTCTGCGAGGAGACGGCGCTGCTCACCCGGCCGGGCGCCCCAGCCGCAGG    339

Qy  100 LysGluGluValAspMetMetLysArgValLeuGluSerLeuAsnLeuAsnIleValGluMet  119
          ||| ||||||  ||| || |||  ||| |||||  |||  :|||||| |::|
Db  340 AAGGAGGTTGAAGCCATGAAGAGAGTACTAGAAAGCCTTAACCTCAACATGGTAGAGATG   399

Qy  120 LysAspGluAsnAlaThrLeuAspGlyGlyAspValPheValLeuPheThrGlyArgGlu---Phe  138
          ||| || ||  ||| ||||| |||||||  |||||| |||||||||  |||
Db  400 GTAGATGAAAATGCAACCTTGGATGGTGGAGACGTCTTATTTACAGGCAGGAGAATTTTTT   459

Qy  139 PheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPhe     158
          |||||||||||  ||| ||   || |||||||||||||||| |||  ||
Db  460 TTTGTAGGCCTTTCAAGGCGAACCAACCAACGTGGGGCAGAAATTCTGGCTGACACATTT   519

Qy  159 LysAspTyrAlaValSerThrValProValHisAspSerLeuHisLeuLysSerPheCys     178
          ||| || ||||||||  ||||| |||  |||  || ||||| || ||||||  ||||
Db  520 AAGGATTATGCTGTCTCCACAGTCCCTGTTCATGATTCTTTGCATCTGAAGAGTTTTTGC   579

Qy  179 SerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGluAlaAlaGlnLysAlaLeu     198
          |||  |||||||  ||  ||||||||||  |:::|||  |  ||| ||||| |||||
Db  580 AGCATGGCTGGACCAAACCTGATTGCTATCGGGTCAAGTGAAGCTGCACAGAAGGCCCTC   639

Qy  199 LysIleMetGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspAlaIle    218
          ||| || ||  |||||| |||||||| ||| |||||| || |||  |||  |
Db  640 AAGACCATGCAACAGATGAGCGACCACCCGTACGACAAGCTGACAGTGCCCGACGATGCC   699
```

```
Qy  219 AlaAlaAsnCysIleTyrLeuAsnIleProSerLysGlyHisValLeuLeuHisArgThr     238
          |||||||||||||||||||||||||||||   |::::|||||| |||||||||||:|
Db  700 GCCGCAAACTGCATCTACTTAAACATTCCAGCAAGGGCCACGTCCTGCTGCACCGAGCC    759

Qy  239 ProGluGluTyrProGluSerAlaLysValPheGluLysLeuLysAspHisMetLeuIle    258
          |||  ||||| |||||  || |||||  ::: |||  ||||||||||| |||  |||
Db  760 CCTGAGGAGTACCCAGAGAGCGCCAAGGTTTTTGAAAAACTGAAGGACCATGCTGATC    819

Qy  259 ProValSerMetSerGluGluGluLysValAspGlyLeuLeuThrCysCysSerValLeu    278
          ||||| ||  :::::  ||| |||  || || |||  ||  ||||||||||||
Db  820 CCCATAGCCAACACAGAACTGGAGAAAGTAGATGGGCACTCACCTGTTGCTGTGCTT    879

Qy  279 IleAsnLys     281
          |||||||
Db  880 ATTAACAAA   888
```

```
RESULT 12
BC056074
LOCUS       BC056074                1344 bp    mRNA    linear   VRT 04-NOV-2003
DEFINITION  Xenopus laevis dimethylarginine dimethylaminohydrolase 1, mRNA
            (cDNA clone MGC:69055 IMAGE:4965611), complete cds.
ACCESSION   BC056074
VERSION     BC056074.1  GI:33417137
KEYWORDS    MGC.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 1344)
  AUTHORS   Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
            and Richardson,P.
  TITLE     Genetic and genomic tools for Xenopus research: The NIH Xenopus
            initiative
  JOURNAL   Dev. Dyn. 225 (4), 384-391 (2002)
  MEDLINE   22341132
   PUBMED   12454917
REFERENCE   2  (bases 1 to 1344)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22338257
   PUBMED   12477932
REFERENCE   3  (bases 1 to 1344)
  AUTHORS   Klein,S. and Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2003) National Institutes of Health, Xenopus Gene
            Collection (XGC), National Institute of Child Health and Human
            Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
            20892-7510, USA
REMARK      NIH-MGC Project
COMMENT     Contact: XGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Igor Dawid
            cDNA Library Preparation: Life Technologies, Inc.
```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 129 Row: j Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

```
FEATURES             Location/Qualifiers
     source          1..1344
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="MGC:69055 IMAGE:4965611"
                     /tissue_type="Spleen, adult Xenopus"
                     /clone_lib="NICHD_XGC_Sp1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1..1344
                     /gene="ddah1-prov"
                     /note="synonym: MGC69055"
                     /db_xref="LocusID:380577"
     CDS             84..938
                     /codon_start=1
                     /product="ddah1-prov protein"
                     /protein_id="AAH56074.1"
                     /db_xref="GI:33417138"
                     /db_xref="LocusID:380577"
                     /translation="MAGIALSGFGRCTHVIVRAIPDSLCQEALRIEKGSVVDISLARK
                     QHEQYVEVLKNKLGLSVIELPADSLPDAPFVEDVAVVCEETALIITRPGAPSRRKETE
                     AVKRALQDLQLNVVEMTDENATLDGGDVLFTGREFFVGLSKRTNQRGAEILADTFKDY
                     AVSTVPVCETLHLKSFCSMAGPNLIAIGSSEAAQKALKTMQQMSDHRYDKLJILPDNAA
                     ANCIYLNIPSKGHVLLHLAPEDFPESAKVFEKLKEHMLFPLPNTELAKVDGCLTCCSI
                     LINKKKSL"
     misc_feature    114..917
                     /note="Amidinotransf; Region: Amidinotransferase. This
                     family contains glycine (EC:2.1.4.1) and inosamine
                     (EC:2.1.4.2) amidinotransferases, enzymes involved in
                     creatine and streptomycin biosynthesis respectively. This
                     family also includes arginine deiminases, EC:3.5.3.6.
                     These enzymes catalyse the reaction: arginine + H2O <==>
                     citrulline + NH3. Also found in this family is the
                     streptococcus anti tumour glycoprotein"
                     /db_xref="CDD:pfam02274"
```

ORIGIN

```
Alignment Scores:
Pred. No.:             8.96e-96
Score:                 1149.50
Percent Similarity:    86.85%      Length:         1344
Best Local Similarity: 77.85%      Matches:         225
Query Match:           79.06%      Conservative:     26
DB:                                Mismatches:       33
                                   Indels:            5
                                   Gaps:              2
```

US-09-889-733B-2 (1-285) x BC056074 (1-1344)

```
Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla  20
         ||||||||||
Db  84 ATGGCTGGACTAGCT--CTGTCAGGCTTTGGGAGATGCACCCACGTTATAGTTCGTGCA  140

Qy  21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAspVal  40
         :||||:||||
Db 141 ATTCCTGACTCTTTATGTCAAGAAGCTCTGAGAGAATAGAGAAAGGCAGTGTGGTGATATC  200

Qy  41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuSerLysGlyLeu  60
         ::||| |||
Db 201 TCCCTGGCTAGGAAACAACATGAGCAGTATGTTGAGGTGCTGAAGAATAAACTGGGGCTG  260

Qy  61 GlnValValGluLeuProAlaAspSerLeuProAspCysValPheValGluAspVal  80
Db 261 AGTGTGATAGAGCTTCCAGCAGATGAATCCCTCCCAGATGCGGCCCTTTGTGGAAGATGTG  320

Qy  81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys  100
Db 321 GCTGTTGTGTGTGAGGAAACTGCCCTGATTACCAGACCTGGGGCACCTAGCACGAGGAAA  380

Qy 101 GlyValAspMetMetLysGlyAlaLeuGlyLysLysGlnLeuAsnIleValGluMetLys  120
Db 381 GGACTGAAGCTGTGAAGAAAGCACTGCAGGATCTTCAGCTAAATGTGGTTGAAATGACA  440

Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal  140
Db 441 GATGAAAATGCTCATTAGTGACGAGAGAGAGAATGTTTTGTTTACAGGAAGGGAATTCTTGTG  500

Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp  160
Db 501 GGTTTGTCGAAAAGGACCAATCACAGTGTGCTGAAATTCTGGCAGATACCTTAAGGAT  560

Qy 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet  180
Db 561 TATGCAGTTTCCACTGTGCCAGTGCGAGACAATTGCACTTGAAGAGTTTCTGCAGCATG  620

Qy 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluAlaAlaGlnLysAlaLeuLysIle  200
Db 621 GCTGGACCAAACCTCATTGCCATTGGCTCGAGTGGAAGCTGCCCAAAAAGCTCTAAAGACC  680

Qy 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspIleAlaAla  220
Db 681 ATGCAGCAGATGAGTGATCACCGGTACGACAAGCTGACCGTGCCTGATAATGCTGCAGCT  740

Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu  240
Db 741 AACTGCATTTATTTAAACATTCCAAGCAAAGGCATGTTCTACTGCATCTTGCCACCTGAA  800

Qy 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal  260
Db 801 GATTTTCCAGAGAGTGCAAAGGTGTTTGAAAACTAAAGGAACACATGCTGTTTCCACTT  860

Qy 261 SerMetSerGluLeuGluLysValAlaAspGlyLeuLeuThrCysCysSerValLeuIleAsn  280
Db 861 CCTAACACCGAGCTGCAAAAGTTGATGATGATGCTGCTCCATCCTTATTAAC  920

Qy 281 LysLys---------ValAspSer  285
Db 921 AAAAAGAGAAAGTCTCGAGTTGATTCA  947
```

RESULT 13

HSM808292

```
LOCUS       HSM808292               3908 bp    mRNA    linear    PRI 30-AUG-2003
DEFINITION  Homo sapiens mRNA; cDNA DKFZp686N2176 (from clone DKFZp86N2176).
ACCESSION   BX648145
VERSION     BX648145.1  GI:34367304
KEYWORDS
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3908)
  AUTHORS   Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  CONSRTM   The German Human cDNA Consortium
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
  COMMENT   Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
```

This clone (DKFZp686N2176) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

```
FEATURES
   source         1..3908
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="DKFZp686N2176"
                  /tissue_type="human endometrium carcinoma cell line"
                  /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
                  DH10B; sites SfiIA + SfiIB"
                  /dev_stage="adult"
ORIGIN
```

Alignment Scores:
Pred. No.:
Score: 4.57e-76   Length: 3908
Matches: 941.00   Matches: 184
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 64.72%   Indels: 0
DB: Gaps: 9

US-09-889-733B-2 (1-285) x HSM808292 (1-3908)

```
Qy  102 ValAspMetLysGluAlaLeuGlyLysGluGlnLeuAsnIleValGluMetLysAsp 121
Db  384 GTTGACATGATGAAAGAAGCATTAGGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGAT 443

Qy  122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGly 141
Db  444 GAAAATGCAACTTTAGATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGGGC 503

Qy  142 LeuSerLysArgThrAsnGlnArgGlyHisLeuLeuAlaGluLeuAspThrArgLysTyr 161
Db  504 CTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTCGCTGATACTTTTAAGGACTAT 563

Qy  162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
Db  564 GCAGTGTCTCCAGTGCCAGTGGCACAGTGGGTTGCATTGAAGAGTTTCTGCAGCATGGCT 623

Qy  182 GlyProAsnLeuIleAlaIleGlyLysSerGluSerAlaGlnLysValAlaLeuLysIleMet 201
Db  624 GGGCCTAACCTGATCGCCATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATG 683

Qy  202 GlnIleMetSerAspHisArgTyrAspLysLeuThrValProAspIleAlaAlaAsn 221
Db  684 CAACAGATGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGCCATAGCAGCAAAC 743

Qy  222 CysIleTyrLeuAsnIleProAsnLysGlyHisValLeuHisValHisArgThrProGluGlu 241
Db  744 TGTATATATCTAAATATCCCCAACAAAGGGACACGTCTTGCTGCACCGACCCCGGAAGAG 803

Qy  242 TyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProValSer 261
Db  804 TATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGC 863

Qy  262 MetSerGluLeuGluLysValAspLysGlyLeuLeuThrCysCysSerValLeuIleAsnLys 281
Db  864 ATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAACAAG 923

Qy  282 LysValAspSer 285
Db  924 AAAGTAGACTCC 935
```

RESULT 14
AK128909
LOCUS   AK128909   2289 bp   mRNA   linear   ROD 09-SEP-2003
DEFINITION   Mus musculus cDNA fis, clone TRACH2023479, highly similar to
             NG,NG-DIMETHYLARGININE DIMETHYLAMINOHYDROLASE 1 (EC 3.5.3.18).
ACCESSION   AK128909

Alignment Scores:
Pred. No.:   2289
Score: 1.54e-73   Length: 2289
Percent Similarity: 910.00   Matches: 176
Best Local Similarity: 98.37%   Conservative: 5
Query Match: 95.65%   Mismatches: 3
DB: 62.59%   Indels: 0
                       10   Gaps: 0

US-09-889-733B-2 (1-285) x AK128909 (1-2289)

```
Qy  102 ValAspMetMetLysGluAlaLeuGlyLeuGlnLeuAsnIleValGluMetLysAsp 121
Db   29 GTTGACATGATGAAAGAAGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGAT 88

Qy  122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGly 141
Db   89 GAAAATGCAACTTTGGATGGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTGTGGGC 148

Qy  142 LeuSerLysArgThrAsnGlnArgGlyHisLeuLeuAlaGluLeuLeuAlaAspThrArgLysTyr 161
Db  149 CTTTCCAAAAGAGACAAATCAACGAGGTGCTGAAATCTTGCTGATACTTTTAAGGACTAC 208

Qy  162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
Db  209 GCAGTCTCCACAGTCCCTGTGGCCGATTCTTTGCATTTAAAGAGTTTTCTGCAGCATGGCC 268

Qy  182 GlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysValAlaLeuLysIleMet 201
Db  269 GGACCCAACCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAGGCCCTCAAGATCATG 328

Qy  202 GlnIleMetSerAspHisArgTyrAspLysLeuThrValProAspIleAlaAlaAsn 221
```

```
Db  329 CAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCGACGACATGCCGGCCAAC 388
            |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Qy  222 CysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGluGlu 241
Db  389 TGCATATATCTAAATATCCCAGCAAAGGGCATGTCTTGCTCACCGAACCCCAGAAGAG 448

Qy  242 TyrProGluSerAlaLysValTyrGluLysLysLysLeuLysAspHisMetLeuIleProValSer 261
Db  449 TACCCAGAAAGCGCAAAGGTCATGAGAAACTCAAGGACCATCTACTGATCCCTGTGAGC 508
Qy  262 MetSerGluLeuGluLysValAspGlyLeuLeuLeuThrCysCysSerValLeuIleAsnLys 281
Db  509 AACTCGGAGATGGAAAAGGTGGACGGCCTGCTCACCTGCTGCTCCGTTTTTATTAACAAG 568

Qy  282 LysValAspSer 285
Db  569 AAGATAGACTCC 580
```

RESULT 15
BC060266
LOCUS   BC060266   3420 bp   mRNA   linear   ROD 25-NOV-2003
DEFINITION Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
            clone IMAGE:5698949), complete cds.
ACCESSION   BC060266
VERSION   BC060266.1 GI:37805213
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 3420)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2 (bases 1 to 3420)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-OCT-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 132 Row: a Column: 19
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis
            This clone has the following problem: The cds is short compared to
            the longest cds in the locus.
FEATURES   Location/Qualifiers
     source   1..3420
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5698949"
                     /tissue_type="Brain, whole, 18.5 dpc mouse embryo"
                     /clone_lib="NIH_BMAP_EQO"
                     /lab_host="DH10B"
                     /note="Vector: pYX-ASC"
     gene   1..3420
                     /gene="Ddah1"
                     /note="synonym: 2510015N06Rik"
                     /db_xref="LocusID:69219"
                     /db_xref="MGI:1916469"
     CDS   360..680
                     /codon_start=1
                     /product="Ddah1 protein"
                     /protein_id="AAH60266.1"
                     /db_xref="GI:37805214"
                     /db_xref="LocusID:69219"
                     /translation="MAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDMAANCI
                     YLNIPSKGHVLLRRTPEBYPESAKVYEKLKDHLLIPVSNSMEMKVDGLLTCCSVFINK
                     KIDS"
     misc_feature   360..659
                     /note="Amidinotransf; Region: Amidinotransferase. This
                     family contains glycine (EC:2.1.4.1) and inosamine
                     (EC:2.1.4.2) amidinotransferases, enzymes involved in
                     creatine and streptomycin biosynthesis respectively. This
                     family also includes arginine deiminases, EC:3.5.3.6.
                     These enzymes catalyse the reaction: arginine + H2O <=>
                     citrulline + NH3. Also found in this family is the
                     streptococcus anti tumour glycoprotein"
                     /db_xref="CDD:pfam02274"
ORIGIN
```

Alignment Scores:
Pred. No.:                                    3420
Score:                   3.31e-59    Length:        3420
Percent Similarity:      756.00      Matches:       145
Best Local Similarity:   97.40%      Conservative:  5
Query Match:             94.16%      Mismatches:    4
DB:                      51.99%      Indels:        0
                         10          Gaps:          0

US-09-889-733B-2 (1-285) x BC060266 (1-3420)

```
Qy  132 LeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAla 151
Db  216 TTGCTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGAACAAATCACCAGAGTGCT 275
Qy  152 GluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrValProValAlaAspGly 171
Db  276 GAAATCTTGGCTGATACTTTTAAGGACTACGCAGTGTCTACAGTCCCTGTGGCCGATTCT 335
Qy  172 LeuHisIleLeuSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSer 191
Db  336 TTGCATTTAAGAGTTTCTCGAGCATGGCCGACCATGGCCGGACCCAACCTGATTGCAATAGGGTCCAGC 395
```

```
Qy  192  GluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLys  211
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  396  GAATCTGCACAGAAGGCCCTCAAGATCATGCAACAGATGAGTGACCATCGTTATGACAAG  455

Qy  212  LeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGly  231
              ||||||||||||||||| :  : :: ||||||||||| | |||||||||
Db  456  CTCACTGTACCCGACGACATGGCCGCCAACTGCATATATCTAAATATCCCCAGCAAAGGG  515

Qy  232  HisValLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLys     251
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  516  CATGTCTTGCTGCACCGAACCCCAGAAGAGTACCCAGAAAGCGCAAAGGTCTATGAGAAA  575

Qy  252  LeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeu  271
              ||||| : :: | ||||||||||||    ||||||| : : |||||||||||||
Db  576  CTCAAGGACCATCTACTGATCCCTGTGAGCAACTCGGAGATGGAAAAGGTGGACGGCTTG  635

Qy  272  LeuThrCysCysSerValLeuIleAsnLysLysValAspSer  285
              ||||||||||| | : : | |||||||||
Db  636  CTCACCTGCTGCTCCGTTTTTATTAACAAGAAGATAGACTCC  677

Search completed: June  8, 2004, 00:22:52
Job time : 3219 secs
```

OM nucleic - nucleic search, using sw model

Copyright GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: June 7, 2004, 20:32:55 ; Search time 3410 Seconds
(without alignments)
10905.667 Million cell updates/sec

Title: US-09-889-733B-1
Perfect score: 858
Sequence: 1 atgccggcctcggccaccc........acaagaggtagactcctga 858

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_pr:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	858	100.0	858	6	BD251986	BD251986 Screen me
2	858	100.0	858	6	AX032835	AX032835 Sequence
3	772	90.0	1624	9	BC033680	BC033680 Homo sapi
4	772	90.0	1633	6	BD094073	BD094073 Shear str
5	772	90.0	1633	9	AB001915	AB001915 Homo sapi
6	772	90.0	4208	6	AX548046	AX548046 Sequence
7	728	84.8	3737	9	BC043235	BC043235 Homo sapi
8	544	63.4	3908	9	HSM808292	BX648145 Homo sapi
9	233	27.2	2895	9	AK098405	AK098405 Homo sapi
10 c	232	27.0	104071	2	AL162260	AL162260 Homo sapi
11 c	232	27.0	158405	2	AC019261	AC019261 Homo sapi
12 c	232	27.0	168953	9	AL360219	AL360219 Human DNA
13 c	148	17.2	83946	9	HSA123M24	AL078459 Human DNA
14	85	9.9	420	6	AX898477	AX898477 Sequence
15	85	9.9	420	6	BD034010	BD034010 Sequence
16 c	79	9.2	110000	2	AC144263_2	Continuation (3 of
17 c	79	9.2	304208	2	AC143028	AC143028 Macaca mu
18	59	6.9	1120	10	BC034505	BC034505 Mus muscu
19	59	6.9	186434	2	AC126412	AC126412 Mus muscu
20	54	6.3	3008	10	AX683109	AX683109 Sequence
21	54	6.3	3008	10	D86041	D86041 Rattus norv
22 c	54	6.3	228940	2	AC126890	AC126890 Rattus no
23	47	5.5	2289	10	AK128909	AK128909 Mus muscu
24	47	5.5	3420	10	BC060266	BC060266 Mus muscu
25	44	5.1	212111	2	AC094482	AC094482 Rattus no
26	43	5.0	626	11	BV064692	BV064692 S212P6014
27 c	43	5.0	186434	2	AC126412	AC126412 Mus muscu
28	43	5.0	200007	2	AC123684	AC123684 Mus muscu
29	26	3.0	1344	5	BC056074	BC056074 Xenopus l
30	25	2.9	58539	2	AC137903	AC137903 Mus muscu
31	23	2.7	2520	10	AF127389	AF127389 Rattus no
32	23	2.7	2526	10	AF301161	AF301161 Mus muscu
33	23	2.7	2526	10	AF301162	AF301162 Mus muscu
34	23	2.7	2771	6	BD233637	BD233637 Nucleic a
35 c	23	2.7	237743	2	AC129831	AC129831 Rattus no
36	22	2.6	1150	5	AB095027	AB095027 Gallus ga
37	22	2.6	90348	1	AF497482	AF497482 Micromono
38	21	2.4	89857	9	AC074097	AC074097 Homo sapi
39 c	21	2.4	98935	2	AC004935	AC004935 Homo sapi
40	21	2.4	162173	2	AC123569	AC123569 Rattus no
41 c	21	2.4	181835	9	AC146227	AC146227 Pan trogl
42 c	21	2.4	210408	2	AC142421	AC142421 Rattus no
43	21	2.4	212026	2	AC146053	AC146053 Pan trogl
44 c	21	2.4	247412	2	AC115249	AC115249 Rattus no
45 c	20	2.3	921	11	BV005803	BV005803 BARC0051

ALIGNMENTS

RESULT 1
BD251986
LOCUS   BD251986   858 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION   Screen method.
ACCESSION   BD251986
VERSION   BD251986.1 GI:33061756
KEYWORDS   JP 202535001-A/1.
SOURCE   Homo sapiens (human).
ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 858)
AUTHORS   Vallance,P.J.T., Leiper,J.M., Whitley,G.S.J. and Charles,I.G.
TITLE   Screen method
JOURNAL   Patent: JP 202535001-A 1 22-OCT-2002;

Pred. No. is the number of results predicted by chance to have a

COMMENT   UNIVERSITY COLLEGE LONDON
OS   Homo sapiens (human)
PN   JP 2002535001-A/1
PD   22-OCT-2002
PF   26-JAN-2000 JP 2000596131
PR   26-JAN-1999 GB   9901705.5,04-JUN-1999 GB   9913066.8 PI
PATRICK JOHN THOMPSON VALLANCE,JAMES MITCHELL LEIPER,GUY ST PI
JOHN WHITLEY,
PI   IAN GEORGE CHARLES
PC   C12N15/09,A01K67/027,A61K31/198,A61K38/00,A61K45/00,A61P3/06,
PC   A61P9/02,
PC   A61P9/04,A61P9/10,A61P9/12,A61P13/12,A61P25/06,A61P25/18, PC
A61P25/28,
PC   A61P29/00,A61P31/04,A61P35/00,C07K16/40,C12N1/15,C12N1/19, PC
C12N1/21,
PC   C12N5/10,C12N9/78,C12Q1/02,C12Q1/34,G01N33/15,G01N33/50 PC
,C12N15/00,C12N5/00,
PC   A61K37/02
CC   Screen method
FH   Key   Location/Qualifiers
FT   CDS   (1)..(858).

FEATURES
    source   1..858
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"

ORIGIN

Query Match   100.0%;   Score 858; DB 6; Length 858;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 858; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGGCCGGCTCGGCCACCCCTCGGCCGCGGCCACCCACGCGTGGTGCGGGCG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGCCGGCTCGGCCACCCCTCGGCCGCGGCCACCCACGCGTGGTGCGGGCG 60

Qy   61 CTACCCGAGTCGCTCTGCCAGCACACGCGCGTGAGAAGGCGCCAAGGGCGGCGAGGAGGTGGACGTC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTACCCGAGTCGCTCTGCCAGCACACGCGCGTGAGAAGGCGCCAAGGGCGGCGAGGAGGTGGACGTC 120

Qy   121 GCCGCGCCGGAACGGCAGCAGCACCAGCTTCACGTGGGCGTGCTGGGCAGCAAGTGGGGCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCCGCGCCGGAACGGCAGCAGCACCAGCTTCACGTGGGCGTGCTGGGCAGCAAGTGGGGCTG 180

Qy   181 CAGGTGGTGGAGCTGCCGGCGACGAGAGCCTTCCGGACTGCCTCTTCGTGGTGGTCGGGAGGACGTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAGGTGGTGGAGCTGCCGGCGACGAGAGCCTTCCGGACTGCCTCTTCGTGGTGGTCGGGAGGACGTG 240

Qy   241 GCCGTGGTGTGCGAGGAGACGGCGCCCTCATCACCCGGGCGGCCCGAGCCCGGAGGAAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GCCGTGGTGTGCGAGGAGACGGCGCCCTCATCACCCGGGCGGCCCGAGCCCGGAGGAAG 300

Qy   301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAAACTTCAGCTCAATATGAGTAGATGATGAAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAAACTTCAGCTCAATATGAGTAGATGATGAAA 360

Qy   361 GATGAAAATGCAACTTAGATGCGGAGAGTGTTTTATTCACAGGCAGAGAATTTTTGTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GATGAAAATGCAACTTTAGATGCGGAGAGTGTTTTATTCACAGGCAGAGAATTTTTGTG 420

Qy   421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480

Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540

Qy   541 GCTGGGCCTAACCTGATCGGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCTGGGCCTAACCTGATCGGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600

Qy   601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660

Db   601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTCGCACCGGAA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTCGCACCGGAA 720
Qy   721 GAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTG 780
Qy   781 AGCATGTCTGAACTGAAAAGGTGGATGGGGCTGCTCACCTGGCTGCTCAGTTTTAATTAAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGCATGTCTGAACTGAAAAGGTGGATGGGGCTGCTCACCTGGCTGCTCAGTTTTAATTAAC 840
Qy   841 AAGAAGGTAGAGACTCCTGA 858
          ||||||||||||||||||||
Db   841 AAGAAGGTAGAGACTCCTGA 858

RESULT 2
AX032835
LOCUS   AX032835   858 bp   DNA   linear   PAT 21-SEP-2000
DEFINITION Sequence 1 from Patent WO0044888.
ACCESSION   AX032835
VERSION   AX032835.1   GI:10279809
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Leiper,J.M., Whitley,G.S., Charles,I.G. and Vallance,P.J.
  TITLE   Screen method
  JOURNAL   Patent: WO 044888-A 1 03-AUG-2000;
            LEIPER JAMES MITCHELL (GB) ; WHITLEY GUY ST JOHN (GB) ; UNIV LONDON
            (GB) ; CHARLES IAN GEORGE (GB) ; VALLANCE PATRICK JOHN THOMPSON
            (GB)
FEATURES   Location/Qualifiers
     source   1..858
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS   1..858
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09947.1"
                     /db_xref="GI:10279810"
                     /db_xref="REMTREMBL:CAC09947"
                     /translation="MAGLGHPSAFGRATHAVVRALPESLCQHAIRSAKGEVDVARAE
                     RQHQLYVGVLGSKLGLGQVVBLPADSLPDCVFVEDVAVVCEETALITRPGAPSRRKEV
                     DMMKEALEKLQLNIVEMKDENATLDGGDVLFTGRRFPVGLSKRTNQRGABILADTFKD
                     YAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVDDI
                     AANCIYLNIPNKGHVLLHRTPEEYPESAKVYEKLKDHMLIPVSMSBLEKVDGLLTCCS
                     VLINKKVDS"
ORIGIN

Query Match   100.0%;   Score 858; DB 6; Length 858;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 858; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGGCCGGCTCGGCCACCCCTCGGCCGCGGCCACCCACGCGTGGTGCGGGCG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGCCGGCTCGGCCACCCCTCGGCCGCGGCCACCCACGCGTGGTGCGGGCG 60
Qy   61 CTACCCGAGTCGCTCTGCCAGCACACGCGCGTGAGAAGGCGCCAAGGGCGGCGAGGAGGTGGACGTC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTACCCGAGTCGCTCTGCCAGCACACGCGCGTGAGAAGGCGCCAAGGGCGGCGAGGAGGAGGTGGACGTC 120
Qy   121 GCCGCGCCGGAACGGCAGCAGCACCAGCTTCACGTGGGCGTGCTGGGCAGCAAGTGGGGCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCCGCGCCGGAACGGCAGCAGCACCAGCTCTCACGTGGGCGTGCTGGGCAGCAAGTGGGGCTG 180

```
Qy  181 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCGTCTTCGTGGAGGACGTG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCGTCTTCGTGGAGGACGTG 240

Qy  241 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCGAGGAAG 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCGAGGAAG 300

Qy  301 GAGGTTGACATGATGAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GAGGTTGACATGATGAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360

Qy  361 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420

Qy  421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480

Qy  481 TATGGAGTCTTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATG 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TATGCAGTCTTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATG 540

Qy  541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGCCCTTAAGATC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGCCCCTTAAGATC 600

Qy  601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660

Qy  661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCGGAA 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCGGAA 720

Qy  721 GAGTATCCAGAAGTGCAAAGTTTATGAGAAACTGAAGGACCATATGTCGATCCCCGTG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGTATCCAGAAGTGCAAAGTTTATGAGAAACTGAAGGACCATATGTCGATCCCCGTG 780

Qy  781 AGCATGTCTGAACTGGAAAAGGTGATGGCTGCTCACCTGCTCAGTTTAATTAAC 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AGCATGTCTGAACTGGAAAAGGTGGATGGCTGCTCACCTGCTCAGTTTAATTAAC 840

Qy  841 AAGAAGGTAGACTCCTGA 858
            ||||||||||||||||||
Db  841 AAGAAGGTAGACTCCTGA 858
```

```
RESULT 3
BC033680
LOCUS        BC033680              1624 bp    mRNA     linear   PRI 07-OCT-2003
DEFINITION   Homo sapiens dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
             clone MGC:45161 IMAGE:5189970), complete cds.
ACCESSION    BC033680
VERSION      BC033680.1  GI:21707414
KEYWORDS     MGC.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 1624)
AUTHORS      Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
             Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
             Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
             Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
             Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
             Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
             Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
             Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
             Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
             McKernan,K.J., Malek,J.A., Gunaraxne,P.H., Richards,S.,
             Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
```

```
             Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
             Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
             Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
             Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
             Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
             Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
             Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE        Generation and initial analysis of more than 15,000 full-length
             human and mouse cDNA sequences
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE      22388257
PUBMED       12477932
REFERENCE    2  (bases 1 to 1624)
AUTHORS      Strausberg,R.
TITLE        Direct Submission
JOURNAL      Submitted (02-JUL-2002) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: National Institutes of Health Intramural
             Sequencing Center (NISC),
             Gaithersburg, Maryland;
             Web site: http://www.nisc.nih.gov/
             Contact: nisc_mgc@nhgri.nih.gov
             Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
             Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
             Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
             Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
             Maduro,Q.L., Masiello,C., Masteri,B., Mastrian,S.D., McCloskey,J.C.,
             McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
             Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
             Young,A., Zhang,L.-H. and Green,E.D.
             Clone distribution: MGC clone distribution information can be found
             through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
             Series: IRAK Plate: 68 Row: m Column: 1
             This clone was selected for full length sequencing because it
             passed the following selection criteria: matched mRNA gi: 6912327.
FEATURES             Location/Qualifiers
     source          1..1624
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:45161 IMAGE:5189970"
                     /tissue_type="Colon, Kidney, Stomach, adult, whole pooled"
                     /clone_lib="NIH_MGC_116"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1..1624
                     /gene="DDAH1"
                     /note="synonym: DDAH"
                     /db_xref="LocusID:23576"
                     /db_xref="MIM:604743"
     CDS             100..957
                     /codon_start=1
                     /product="DDAH1 protein"
                     /protein_id="AAH33680.1"
                     /db_xref="GI:21707415"
                     /db_xref="LocusID:23576"
                     /translation="MAGLGHPAAFGRATHAVVRALPESLGQHALRSAKGEEVVARAE
                     RQHQLYVGVLGSKLGLGQVVELPADESLPDCVFVEDVAVVCEETALITRPGAPSRRKEV
                     DMMKKEALEKLQLNIVEMKDENATLDQGEILADTFKD
                     YAVSTVPVADGLHLKSPCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDI
                     AANCIYLNIPNKDHVLLHRTPEYPESAKVYEKLKDEHMLIPVSMSELKVDGLLTCCS
                     VLINKKVDS"
     misc_feature    133..936
                     /note="Amidinotransf; Region: Amidinotransferase. This
```

```
LOCUS        BD094073                 1633 bp    DNA    linear    PAT 27-AUG-2002
DEFINITION   Shear stress-responsive DNAs.
ACCESSION    BD094073
VERSION      BD094073.1  GI:22639661
KEYWORDS     WO 0125427-A/34.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 1633)
AUTHORS      Nojima,H., Yoshisue,H., Obayashi,M., Ota,T., Kawabata,A.,
             Sakurada,K., Kuga,T., Sekine,S., Nakamura,Y. and Sugan,S.
TITLE        Shear stress-responsive DNAs
JOURNAL      Patent: WO 0125427-A 34 12-APR-2001;
             KYOWA HAKKO KOGYO CO LTD,HIROSHI NOJIMA,HAJIME YOSHISUE, MASAYA
             OBAYASHI, TOSHIO OTA,AYAKO KAWABATA,KAZUHIRO SAKURADA,TETSURO KUGA,
             SUSUMU SEKINE, YUSUKE NAKAMURA,SUMIO SUGANO
COMMENT      OS  Homo sapiens (human)
             PN  WO 0125427-A/34
             PD  12-APR-2001
             PF  02-OCT-2000 WO 2000JP006840
             PR  01-OCT-1999 JP 99P  280976
             PI  HIROSHI NOJIMA,HAJIME YOSHISUE,MASAYA OBAYASHI,TOSHIO OTA, PI
             PI  AYAKO KAWABATA,
             PI  KAZUHIRO SAKURADA,TETSURO KUGA,SUSUMU SEKINE,YUSUKE NAKAMURA,
             PI  SUMIO SUGANO
             PC  C12N15/12,C07K14/435,C07K16/18,C12P21/02,C12Q1/68,A61K38/00,
             PC  A61K39/395,
             PC  A61K48/00,A61P9/10,A61P9/10,G01N33/50,G01N33/53
             CC
             FH  Key             Location/Qualifiers
             FT  CDS             (323)..(1177).
FEATURES     Location/Qualifiers
     source  1..1633
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
ORIGIN
```

Query Match   90.0%; Score 772; DB 6; Length 1633;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy     23  CGGCCTTCGGCGCGGCCACCACCGCCGTGGTGCGGCGCTACCCGAGTCGCTCTGCCAGC   82
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    345  CCGCCTTCGGCGCGGCCACCACCGCCGTGGTGCGGCGCTACCCGAGTCGCTCTGCCAGC  404

Qy     83  ACGCGCTGAGAAGCGGCCAAGGGCGAGGAGGTGGACGTGGCTGGCAAGCGGCAGCACC  142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    405  ACGCGCTGAGAAGCGGCCAAGGGCGAGGAGGTGGACGTGGCTGGCAAGCGGCAGCACC  464

Qy    143  AGCTCTACGTGGGCGTGCTGGCAGCAAGCTGGGCTGGTGCGAGCTGCCGGCCG  202
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    465  AGCTCTACGTGGGCGTGCTGGCAGCAAGCTGGGCTGGTGCGAGCTGCCGGCCG  524

Qy    203  ACGAGAGCCTTCCGGACTGCGTCTTCGTGGCAGACTGGCCGTGGTGTGCGAGGAGACGG  262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    525  ACGAGAGCCTTCCGGACTGCGTCTTCGTGGCAGACTGGCCGTGGTGTGCGAGGAGACGG  584

Qy    263  CCCTCATCACCCGACCCGGGGGCGCCGAGCCCGGAGGAAGGAGGTTGACATGATGAAAGAAG  322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    585  CCCTCATCACCCGACCCGGGGGCGCCGAGCCCGGAGGAAGGAGGTTGACATGATGAAAGAAG  644

Qy    323  CATTAGAAAAAACTTCAGCTCAATATAGTAGTAGACTTTAGATG  382
           ||||||||||||||||||||||||||||||||||||||||||||
Db    645  CATTAGAAAAAACTTCAGCTCAATATAGTAGTAGACTTTAGATG  704

Qy    383  GCGGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC  442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    705  GCGGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC  764

Qy    443  AACGAGGTGCTGAAATCTTGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAG  502
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

(left panel)

```
ORIGIN
```
family contains glycine (EC:2.1.4.1) and inosamine
(EC:2.1.4.2) amidinotransferases, enzymes involved in
creatine and streptomycin biosynthesis respectively. This
family also includes arginine deiminases, EC:3.5.3.6.
These enzymes catalyse the reaction: arginine + H2O <=>
citrulline + NH3. Also found in this family is the
Streptococcus anti tumor glycoprotein"
/db_xref="CDD:pfam02274"

Query Match   90.0%; Score 772; DB 9; Length 1624;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy     23  CGGCCTTCGGCGCGGCCACCACCGCCGTGGTGCGGCGCTACCCGAGTCGCTCTGCCAGC   82
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122  CCGCCTTCGGCGCGGCCACCACCGCCGTGGTGCGGCGCTACCCGAGTCGCTCTGCCAGC  181

Qy     83  ACGCGCTGAGAAGCGGCCAAGGGCGAGGAGGTGGACGTGGCTGGCAAGCGGCAGCACC  142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182  ACGCGCTGAGAAGCGGCCAAGGGCGAGGAGGTGGACGTGGCTGGCAAGCGGCAGCACC  241

Qy    143  AGCTCTACGTGGGCGTGCTGGCAGCAAGCTGGGCTGGTGCGAGCTGCCGGCCG  202
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242  AGCTCTACGTGGGCGTGCTGGCAGCAAGCTGGGCTGGTGCGAGCTGCCGGCCG  301

Qy    203  ACGAGAGCCTTCCGGACTGCGTCTTCGTGGCAGACTGGCCGTGGTGTGCGAGGAGACGG  262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302  ACGAGAGCCTTCCGGACTGCGTCTTCGTGGCAGACTGGCCGTGGTGTGCGAGGAGACGG  361

Qy    263  CCCTCATCACCCGACCCGGGGGCGCCGAGCCCGGAGGAAGGAGGTTGACATGATGAAAGAAG  322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362  CCCTCATCACCCGACCCGGGGGCGCCGAGCCCGGAGGAAGGAGGTTGACATGATGAAAGAAG  421

Qy    323  CATTAGAAAAAACTTCAGCTCAATATAGTAGTAGACTTTAGATG  382
           ||||||||||||||||||||||||||||||||||||||||||||
Db    422  CATTAGAAAAAACTTCAGCTCAATATAGTAGTAGACTTTAGATG  481

Qy    383  GCGGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC  442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482  GCGGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC  541

Qy    443  AACGAGGTGCTGAAATCTTGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAG  502
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542  AACGAGGTGCTGAAATCTTGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAG  601

Qy    503  TGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAA  562
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602  TGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAA  661

Qy    563  TTGGGTCTCAGTGAATCTGCCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACC  622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662  TTGGGTCTCAGTGAATCTGCCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACC  721

Qy    623  GCTACGACAAAACTCACTGTGCCTGATGACATAGCACGCAAACTGTATATATCCAGAAAGTGCAAAGG  682
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    722  GCTACGACAAAACTCACTGTGCCTGATGACATAGCACGCAAACTGTATATATCCAGAAAGTGCAAAGG  781

Qy    683  CCAACAAAGGCACGTCTTGCTGCCGAAGGAGTATCCCGGAAGCATGTCTGAACTGGAAAGG  742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    782  CCAACAAAGGGCACGTCTTGCTGCCGAAGGAGTATCCCGGAAGCATGTCTGAACTGGAAAGG  841

Qy    743  TTTATGAGAAACTGAAGGACCATATGCTGAGCATGTCCTGAACTCGGAAAAGG  802
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    842  TTTATGAGAAACTGAAGGACCATATGCTGAGCATGTCCTGAACTCGGAAAAGG  901

Qy    803  TGGATGGGCTGCTCACCTGCTTCAGTTTTAATTAACAAGAA  845
           |||||||||||||||||||||||||||||||||||||||||
Db    902  TGGATGGGCTGCTCACCTGCTTCAGTTTTAATTAACAAGAA  944
```

RESULT 4
BD094073

ORIGIN

```
Query Match      90.0%;  Score 772;  DB 9;  Length 1633;
Best Local Similarity  99.9%;  Pred. No. 0;
Matches  822;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy       23 CCGCCTTCGGCCGGCCACCCACGCCGTGGTGCGGGCCCTACCCGAGTCGCTCTGCCAGC    82
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      345 CCGCCTTCGGCCGGCCACCCACGCCGTGGTGCGGGCCCTACCCGAGTCGCTCTGCCAGC   404

Qy       83 ACGGGCTGAGAAGGCCAAGGGACGTGCCGAGGGGCGTGCCGCCCGCGGAACGGCAGCACC   142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      405 ACGCGCTGAGAAGGCCAAGGGACGTGCCGAGGGGCGTGCCGCCCGCGGAACGGCAGCACC   464

Qy      143 AGTCTACGTGGGCGTGCTGGCAGCAAGCAAGGTGGTCGAGTGCCGGCCG   202
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      465 AGTCTACGTGGGCGTGCTGGCAGCAAGCAAGGTGGTCGAGTGCCGGCCG   524

Qy      203 ACGAGAGCCTTCGGACTGCGTCTTCGTGGAGACGTGCCGTGGTGCGGGACGG   262
            ||||||||| |||||| |||||||||||||||||||||||||||||||||||
Db      525 ACGAGAGCCTTCGGACTGCGTCTTCGTGGAGACGTGCCGTGGTGCGGGACGG   584

Qy      263 CCCTCATCACCCGACCCGGGGCGCCCGAGCCCGAGCCGGAGGAGGAGGTTGAACATGATGAAAGAAG   322
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      585 CCCTCATCACCCGACCCGGGGCGCCCGAGCCCGAGCCGGAGGAGGAGGTTGAACATGATGAAAGAAG   644

Qy      323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGAGATGAAAGATGAAAGATGAAAAGATGAAACTTTAGATG   382
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      645 CATTAGAAAAACTTCAGCTCAATATAGTAGAGAGATGAAAGATGAAAGATGAAAAGATGAAACTTTAGATG   704

Qy      383 GCCGAGATGTTTATTCACAGGCAGAGAATTTTGTGGCCTTCCAAAAGACAAATC   442
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      705 GCGAGATGTTTATTCACAGGCAGAGAATTTTGTGGCCTTCCAAAAGACAAATC   764

Qy      443 AACGAGGTGCTGAAATCTTGGCTGATACTTTGGCTGATACTTTTAAGGACTATGCCAGTCTCCACAGTGCCAG   502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      765 AACGAGGTGCTGAAATCTTGGCTGATACTTTGGCTGATACTTTTAAGGACTATGCCAGTCTCCACAGTGCCAG   824

Qy      503 TGGCAGATGGGTTGCATTTGAAGAGTTCTGCAGCATGTCGGGCCTAACCTGATCGCAA   562
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      825 TGGCAGATGGGTTGCATTTGAAGAGTTCTGCAGCATGTCGGGCCTAACCTGATCGCAA   884

Qy      563 TTGGGTCTCAGTGAATCTGCACAGAAGGCCCTTAGATCATGCAAGAAGATGAGTGACCACC   622
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      885 TTGGGTCTCAGTGAATCTGCACAGAAGGCCCTTAGATCATGCAAGAAGATGAGTGACCACC   944

Qy      623 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC   682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      945 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC  1004

Qy      683 CCAACAAAGGGCACGTCTTCGTGCCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAAGG   742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1005 CCAACAAAGGGCACGTCTTCGTGCCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAAGG  1064

Qy      743 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGG   802
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1065 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGG  1124

Qy      803 TGGATGGGCTGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA   845
            |||||||||||||||||||||||||||||||||||||||||||
Db     1125 TGGATGGGCTGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA  1167
```

RESULT 6

```
AX548046
LOCUS       AX548046                4208 bp    DNA     linear   PAT 26-NOV-2002
DEFINITION  Sequence 22 from Patent WO02066654.
ACCESSION   AX548046
VERSION     AX548046.1  GI:25813142
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Db      765 AACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCCAGTCTCCACAGTGCCAG   824

Qy      503 TGGCAGATGGGTTGCATTTGAAGAGTTCTGCAGCATGTCGGGCCTAACCTGATCGCAA   562
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      825 TGGCAGATGGGTTGCATTTGAAGAGTTCTGCAGCATGTCGGGCCTAACCTGATCGCAA   884

Qy      563 TTGGGTCTCAGTGAATCTGCACAGAAGGCCCTTAGATCATGCAACAGATGAGTGACCACC   622
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      885 TTGGGTCTCAGTGAATCTGCACAGAAGGCCCTTAGATCATGCAACAGATGAGTGACCACC   944

Qy      623 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC   682
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      945 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC  1004

Qy      683 CCAACAAAGGGCACGTCTTCGTGCCACCGAACCCCGGAAGAGTATCCCGAAGAGTGCAAAGG   742
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1005 CCAACAAAGGGCACGTCTTCGTGCCACCGAACCCCGGAAGAGTATCCCGAAGAGTGCAAAGG  1064

Qy      743 TTTATGAGAAACTGAAGGACATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGG   802
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1065 TTTATGAGAAACTGAAGGACATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGG  1124

Qy      803 TGGATGGGCTGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA   845
            |||||||||||||||||||||||||||||||||||||||||||
Db     1125 TGGATGGGCTGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA  1167
```

RESULT 5

```
AB001915
LOCUS       AB001915                1633 bp    mRNA    linear   PRI 21-SEP-2000
DEFINITION  Homo sapiens mRNA for NG,NG-dimethylarginine
            dimethylaminohydrolase, complete cds.
ACCESSION   AB001915
VERSION     AB001915.1  GI:4160665
KEYWORDS    NG,NG-dimethylarginine dimethylaminohydrolase.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kimoto,M., Miyatake,S., Sasagawa,T., Yamashita,H., Okita,M.,
            Oka,T., Ogawa,T. and Tsuji,H.
  TITLE     Purification, cDNA cloning and expression of human
            NG,NG-dimethylarginine dimethylaminohydrolase
  JOURNAL   Eur. J. Biochem. 258 (2), 863-868 (1998)
   MEDLINE  99089653
   PUBMED   9874257
REFERENCE   2  (bases 1 to 1633)
  AUTHORS   Kimoto,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-1997) Masumi Kimoto, Okayama Prefectural
            University, Faculty of Health and Welfare Science, Department of
            Nutritional Science; Kuboki 111, Soja, Okayama 719-1197, Japan
            (E-mail:kimoto@fhw.oka.-pu.ac.jp, Tel:0866-94-2158,
            Fax:0866-94-2158)
FEATURES             Location/Qualifiers
     source          1..1633
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="Kidney"
     CDS             323..1180
                     /EC_number="3.5.3.18"
                     /codon_start=1
                     /product="NG,NG-dimethylarginine dimethylaminohydrolase"
                     /protein_id="BAA77117.1"
                     /db_xref="GI:4160666"
                     /translation="MAGLGHPAAFGRATHAVVRALPESLGQHALRSAKGEVDVARAE
                     RQHPLYVGLGSKLGLQVVELPADESLPDCVVEDVAVVCEETALITRPGAPSRRKEV
                     DMMKEALEKLQLNIVEMKDENATLDGDGVLFTGREFFVGLSKRTNQRGAEILADTFKD
                     YAVSTVPVADGLHLKSPCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDI
                     AANCIYLNIPNKGHVLLHRTPEYPESAKVYEKLKDHMLIPVSMSELEKVDGLLTCCS
                     VLINKKVDS"
```

REFERENCE 1
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Astromoff,A., Au-Young,J., Baughn,M.R., Ding,L., Duggan,B.M.,
          Forsythe,I.J., Gietzen,K.J., Griffin,J.A., Lee,E.A., Lu,Y.,
          Richardson,T.W., Ring,H.Z., Sanjanwala,M.M., Swarnakar,A.,
          Walia,N.K., Warren,B.A., Xu,Y., Yue,H. and Zebarjadian,Y.
TITLE   Drug metabolizing enzymes
JOURNAL   Patent: WO 02066654-A 22 29-AUG-2002;
          Incyte Genomics,Inc. (US)
FEATURES source
          Location/Qualifiers
          1..4208
          /organism="Homo sapiens"
          /mol_type="unassigned DNA"
          /db_xref="taxon:9606"
          /note="Incyte ID No: 7671089CB1"

ORIGIN

Query Match   90.0%; Score 772; DB 6; Length 4208;
Best Local Similarity   99.9%;   Pred. No. 0;
Matches 822;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

Qy   23 CCGCCTTCGGCGCGGCCACCCCACCCAGCGCCGTGGTTCGCGGGGCGCTACCCGAGTTCGGCTCGCTCGGCCAGC 82
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   390 CCGCCTTCGGCGCGGCCACCCCACCCAGCGCCGTGGTTCGCGGGGCGCTACCCGAGTTCGGCTCGCTCGGCCAGC 449

Qy   83 ACCGCGTGAGAAGCGCCAAGGGCGAGGAGGTGACGTCGCCCGCGCCGGAACGGCAGCACC 142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   450 ACCGCGTGAGAAGCGCCAAGGGCGAGGAGGTGACGTCGCCCGCGCCGGAACGGCAGCACC 509

Qy   143 AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGTGGTGGAGCTGCCGGCCG 202
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   510 AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGTGGTGGAGCTGCCGGCCG 569

Qy   203 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGACGACGTGGCCGTGGTGTGCGAGGAGACGG 262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   570 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGACGACGTGGCCGTGGTGTGCGAGGAGACGG 629

Qy   263 CCCTCATCACCCGACCCCGGCGCCGGGGCGGCCGGAGCCGGAGGAGGAGGTTGACATGATGAAAGAAG 322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   630 CCCTCATCACCCGACCCCGGCGCCGGGGCGGCCGGAGCCGGAGGAGGAGGTTGACATGATGAAAGAAG 689

Qy   323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAATGCAACTTTAGATG 382
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   690 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAATGCAACTTTAGATG 749

Qy   383 GCGGAGATGTTTATTCACAGGCAGAGAATTTTGTGGGCCTTTCCAAAAGGACAAATC 442
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   750 GCGGAGATGTTTATTCACAGGCAGAGAATTTTGTGGGCCTTTCCAAAAGGACAAATC 809

Qy   443 AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAG 502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   810 AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAG 869

Qy   503 TGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGCATGGCCTAACCTGATCGCAA 562
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   870 TGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGCATGGCCTAACCTGATCGCAA 929

Qy   563 TTGGGTCTCAGTGAATCTGCACAGAAGGCCCTAAGATCATCGAAGACCCTAAGATCATCG 622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   930 TTGGGTCTCAGTGAATCTGCACAGAAGGCCCTAAGATCATCGAAGACCCTAAGATCATCG 989

Qy   623 GCTACGGACAAAACTCACTGTGCCTGATGACATAGCAGCAGGAAACTGTATATATCCTAAATATCC 682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   990 GCTACGGACAAAACTCACTGTGCCTGATGACATAGCAGCAGGAAACTGTATATATCCTAAATATCC 1049

Qy   683 CCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAAGAGTATCCGAAGAGTGCAAAGG 742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 CCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAAGAGTATCCGAAGAGTGCAAAGG 1109

Qy   743 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCGAACTGGAAAAAGG 802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCGAACTGGAAAAAGG 1169

Qy   803 TGGATGGGCTGCTCACCTGCTCCAGTTTAATTAACAAGAA 845
            ||||||||||||||||||||||||||||||||||||||||
Db   1170 TGGATGGGCTGCTCACCTGCTCCAGTTTAATTAACAAGAA 1212

RESULT 7
BC043235
LOCUS   BC043235   3737 bp   mRNA   linear   PRI 30-SEP-2003
DEFINITION Homo sapiens dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
            clone IMAGE:5295723), partial cds.
ACCESSION   BC043235
VERSION   BC043235.2   GI:34783628
KEYWORDS   .
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 3737)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22388257
PUBMED   12477932
REFERENCE   2 (bases 1 to 3737)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (09-JAN-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   On Sep 16, 2003 this sequence version replaced gi:28175756.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 76 Row: j Column: 15.
FEATURES   Location/Qualifiers
source   1..3737

```
Db    421 ATGGCTGGGCCTAACCTGATCGCAATTGGGTCTCAGTGAATCTGCACAGAAGGCCCTTAAG  480

Qy    598 ATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCA  657
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCA  540

Qy    658 GCAAACTGTATATATCTAAATATCCCCAACAAACAAAGGGCACAGTCTTGCTGCACCGAACCCCG  717
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCAAACTGTATATATCTAAATATCCCCAACAAACAAAGGGCACAGTCTTGCTGCACCGAACCCCG  600

Qy    718 GAAGAGTATCCAGAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGTGATCCCC  777
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GAAGAGTATCCAGAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGTGATCCCC  660

Qy    778 GTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATT  837
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTTAATT  720

Qy    838 AACAAGAA  845
          ||| ||||
Db    721 AACACGAGAA  728
```

```
RESULT 8
HSM808292
LOCUS       HSM808292               3908 bp    mRNA            linear   PRI 30-AUG-2003
DEFINITION  Homo sapiens mRNA; cDNA DKFZp686N2176 (from clone DKFZp686N2176).
ACCESSION   BX648145
VERSION     BX648145.1   GI:34367304
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3908)
  AUTHORS   Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  CONSRTM   The German Human cDNA Consortium
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
  COMMENT   Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            This clone (DKFZp686N2176) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1..3908
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686N2176"
                     /tissue_type="human endometrium carcinoma cell line"
                     /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
                     DH10B; sites SfiIA + SfiIB"
                     /dev_stage="adult"
ORIGIN
```

```
Query Match             63.4%;   Score 544;  DB 9; Length 3908;
Best Local Similarity  100.0%;   Pred. No. 2.9e-283;
Matches 544;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    302 AGGTTGACATGATGAAAGAAGAAGCATTAGAAAAACTTCAGCTCAATATATAGTAGAGATGAAAG  361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 AGGTTGACATGATGAAAGAAGAAGCATTAGAAAAACTTCAGCTCAATATATAGTAGAGATGAAAG  441

Qy    362 ATGAAAATGCAACTTTAGATGGGCGGAGAGTGTTTATTCACAGGCAGGAGAATTTTTGTGG  421
```

Left column (rotated feature listing):

```
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /clone="IMAGE:5295723"
                     /tissue_type="Testis"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
     gene            <1..3737
                     /gene="DDAH1"
                     /note="synonym: DDAH"
                     /db_xref="LocusID:23576"
                     /db_xref="MIM:604743"
     CDS             <1..741
                     /gene="DDAH1"
                     /codon_start=1
                     /product="DDAH1 protein"
                     /protein_id="AAH43235.2"
                     /db_xref="GI:34783629"
                     /db_xref="LocusID:23576"
                     /db_xref="MIM:604743"
                     /translation="VARAERQHQLYVGVLGSKLGLQGVVELPADESLPDCVFVEDVAVV
                     CEETALTTRPGAPSRRKEVDMVMKEALEKLQLNIVEMKDENATLDGGDVLFTGREEFVG
                     LSKRTNQRGAEILADTFKDYAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALK
                     IMQQMSDHRYDKLTVPDDIAANCIYLNIPNKGHVLLHRTPEYPESAKVYEKLKDHML
                     IPVSMSELEKVDGLLTCCSVLINKKVDS"
                     /db_xref="CDD:pfam02274"
     misc_feature    1..720
                     /gene="DDAH1"
                     /note="Amidinotransf; Region: Amidinotransferase. This
                     family contains glycine (EC:2.1.4.1) and inosamine
                     (EC:2.1.4.2) amidinotransferases, enzymes involved in
                     creatine and streptomycin biosynthesis respectively. This
                     family also includes arginine deiminases, EC:3.5.6.
                     These enzymes catalyse the reaction: arginine + H2O <=>
                     citrulline + NH3. Also found in this family is the
                     Streptococcus anti tumor glycoprotein"
ORIGIN
```

```
Query Match             84.8%;   Score 728;   DB 9;  Length 3737;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 728;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    118 GTCGCCCGCCGGAACGGCCAGCACCAGCTCTACGTGGGGGTGCTGGGCAGCAAGCTGGGG  177
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCGCCCGCCGGAACGGCCAGCACCAGCTCTACGTGGGGGTGCTGGGCAGCAAGCTGGGG  60

Qy    178 CTGCAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTTGGAGGAC  237
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTTGGAGGAC  120

Qy    238 GTGGCCGTGGTGTGTGGCGAGGAGACCGCCCTCATCACCCGACACCGACCGGGGCGCCGAGGG  297
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGGCCGTGGTGTGTGGCGAGGAGACCGCCCTCATCACCCGACACCGACCGGGGCGCCGAGGG  180

Qy    298 AAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATG  357
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATG  240

Qy    358 AAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGCAGAATTTTT  417
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGCAGAATTTTT  300

Qy    418 GTGGGCCTTTCCAAAAGGACAAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAG  477
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GTGGGCCTTTCCAAAAGGACAAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAG  360

Qy    478 GACTATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTGAAGAGTTTCTGCAGC  537
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GACTATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTGAAGAGTTTCTGCAGC  420

Qy    538 ATGGCTGGGCCTAACCTGATCGCAATTGGGTCTCAGTGAATCTGCACAGAAGGCCCTTAAG  597
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   442 ATGAAAAATGCAACTTTAGATGCGGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGG 501
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   422 GCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACT 481
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   502 GCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACT 561

Qy   482 ATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATGG 541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   562 ATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATGG 621

Qy   542 CTGGGCCTAACCTGACTGATCGCAATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCA 601
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   622 CTGGGCCTAACCTGACTGATCGCAATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCA 681

Qy   602 TGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGCCTGATGACATAGCAGCAA 661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682 TGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGCCTGATGACATAGCAGCAA 741

Qy   662 ACTGTATATATCTAAATATCCCAACAAAAGGCACGTCTTCGTGCACCGAACCCCGGAAG 721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   742 ACTGTATATATCTAAATATCCCAACAAAAGGGCACGTCTTCGTGCACCGAACCCCGGAAG 801

Qy   722 AGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCACCATATGCTACTCCCGTGA 781
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   802 AGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCACCATATGCTACTCCCGTGA 861

Qy   782 GCATGTCTGAACTGGAAAAGGTGGATGGGCGTCCTACCTGCTGCTCCAGTTTTAATTAACA 841
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   862 GCATGTCTGAACTGGAAAAGGTGGATGGGCGTCCTACCTGCTGCTCCAGTTTTAATTAACA 921

Qy   842 AGAA 845
             ||||
Db   922 AGAA 925
```

RESULT 9
AK098405
LOCUS   AK098405   2895 bp   mRNA   linear   PRI 12-SEP-2003
DEFINITION Homo sapiens cDNA FLJ25539 fis, clone CBR09481.
ACCESSION   AK098405
VERSION   AK098405.1 GI:21758410
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
           Fukuzumi,Y., Fujimori,Y., Komiyama,M., Suzuki,Y., Hata,H.,
           Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T.,
           Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A.,
           Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.
TITLE   NEDO human cDNA sequencing project
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 2895)
AUTHORS   Sugano,S. and Suzuki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
           University of Tokyo, Laboratory of Genome Structure, Human Genome
           Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
           (E-mail:fldna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
           Fax:81-3-5449-5416)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
           Economy, Trade and Industry of Japan; cDNA full insert sequencing:
           Research Association for Biotechnology (RAB); cDNA library
           construction and 5'-end one pass sequencing: Institute of Medical
           Science, University of Tokyo, Laboratory of Genome Structure, Human
           Genome Center; 3'-end one pass sequencing: RAB; clone selection for
           full insert sequencing: RAB and Helix Research Institute.
FEATURES   Location/Qualifiers
     source   1..2895
                     /organism="Homo sapiens"

                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CBR09481"
                     /tissue_type="brain"
                     /clone_lib="CBR"
                     /note="cloning vector: pME18SFL3"
ORIGIN

```
Query Match          27.2%;  Score 233;  DB 9;  Length 2895;
Best Local Similarity 100.0%;  Pred. No. 1.7e-114;
Matches  233;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   412 TTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACT 471
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 TTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACT 486

Qy   472 TTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTGAAGAGTTTC 531
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 TTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTGAAGAGTTTC 546

Qy   532 TGCAGCATGGCTGGGCCTAACCTGACTGATCGCAATTGGGTCTAGTGAATCTGCACAGAGGCC 591
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 TGCAGCATGGCTGGGCCTAACCTGACTGATCGCAATTGGGTCTAGTGAATCTGCACAGAGGCC 606

Qy   592 CTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCC 644
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 CTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCC 659
```

RESULT 10
AL162260/c
LOCUS   AL162260   104071 bp   DNA   linear   HTG 10-JUL-2001
DEFINITION Homo sapiens chromosome 1 clone RP4-611J7 map p22.3-31.1, 8
            unordered pieces.
ACCESSION   AL162260
VERSION   AL162260.4 GI:9863624
KEYWORDS   HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   McLay,K.
TITLE   Direct Submission
JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
           CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
           requests: clonerequest@sanger.ac.uk
COMMENT   On Aug 21, 2000 this sequence version replaced gi:9212914.
           ------------------- Genome Center:
           Center: Sanger Centre
           Center code: SC
           Web site: http://www.sanger.ac.uk
           Contact: humquery@sanger.ac.uk
           ------------------- Project Information
           Center project name: dr611J7
           ------------------- Summary Statistics
           Assembly program: XGAP4; version 4.5
           Sequencing vector: plasmid; L08752; 100% of reads
           Chemistry: Dye-terminator ET-amersham; 16% of reads Chemistry:
           Dye-terminator Big Dye; 83% of reads
           Consensus quality: 100691 bases at least Q40
           Consensus quality: 101774 bases at least Q30
           Consensus quality: 102472 bases at least Q20
           Insert size: 10337l; sum-of-contigs
           Insert size: 112648; 4.2% error; agarose-fp
           Quality coverage: 4.76x in Q20 bases; sum-of-contigs Quality
           coverage: 4.41x in Q20 bases; agarose-fp
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 8 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.

```
*  This record will be updated with the finished sequence
*  as soon as it is available and the accession number will
*  be preserved.
*         1    4008: contig of 4008 bp in length
*      4009    4108: gap of 100 bp
*      4109   17620: contig of 13512 bp in length
*     17621   17720: gap of 100 bp
*     17721   29332: contig of 11612 bp in length
*     29333   29432: gap of 100 bp
*     29433   34595: contig of 5163 bp in length
*     34596   34695: gap of 100 bp
*     34696   51023: contig of 16328 bp in length
*     51024   51123: gap of 100 bp
*     51124   56471: contig of 5348 bp in length
*     56472   56571: gap of 100 bp
*     56572   62357: contig of 5786 bp in length
*     62358   62457: gap of 100 bp
*     62458  104071: contig of 41614 bp in length.

FEATURES           Location/Qualifiers
     source        1..104071
                   /organism="Homo sapiens"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:9606"
                   /chromosome="1"
                   /map="P22.3-31.1"
                   /clone="RP4-611J7"
                   /clone_lib="RPCI-4"
     misc_feature  1..4008
                   /note="assembly_fragment:00515
                   fragment_chain:1
                   clone_end:SP6
                   vector_side:left"
     misc_feature  4109..17620
                   /note="assembly_fragment:00727
                   fragment_chain:1"
     misc_feature  17721..29332
                   /note="assembly_fragment:00262
                   fragment_chain:1"
     misc_feature  29433..34595
                   /note="assembly_fragment:00257
                   fragment_chain:1"
     misc_feature  34696..51023
                   /note="assembly_fragment:00672
                   fragment_chain:1"
     misc_feature  51124..56471
                   /note="assembly_fragment:00205"
     misc_feature  56572..62357
                   /note="assembly_fragment:01056"
     misc_feature  62458..104071
                   /note="assembly_fragment:00375
                   clone_end:T7
                   vector_side:right"

ORIGIN

Query Match           27.0%;  Score 232;  DB 2;  Length 104071;
Best Local Similarity 99.6%;  Pred. No. 5.3e-114;
Matches 282;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     23 CCGCCTTCGGCCGGCCCACCCAGGCCGTGTGTGCGGGCGCTACCCGAGTCGCTTCTGCCAGC     82
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  37550 CCGCCTTCGGCCGGCCCACCCAGGCCGTGTGTGCGGGCGCTACCCGAGTCGCTTCTGCCAGC  37491

Qy     83 ACGCGCTGAGAAGCGCCAAGGGCGGAGGTGGACGTGCGCCCGCGGCCGGAACGGCAGCACC    142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  37490 ACGCGCTGAGAAGCGCCAAGGGCGGAGGTGGACGTGCGCCCGCGGCCGGAACGGCAGCACC  37431

Qy    143 AGCTCTACGTGGGCGGTGCTGGGCGCAGCAAGCTGGGCGGTGCTGGGCGGTGTGGAGGTGGCCG    202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  37430 AGCTCTACGTGGGCGGTGCTGGGCGCAGCAAGCTGGGCGGTGCTGGGCGGTGTGGAGGTGGCCG  37371

Qy    203 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGGACGG    262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  37370 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGGACGG  37311

Qy    263 CCCTCATCACCCGACCCGGGGGCGGGGCCCGAGCCCGAGCCGAGGAGGAGGT    305
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  37310 CCCTCATCACCCGACCCGGGGGCGGGGCCCGAGCCCGAGCCGAGGAGGAGGT  37268
```

```
RESULT 11
AC019261
LOCUS       AC019261              158405 bp    DNA     linear   HTG 12-MAR-2000
DEFINITION  Homo sapiens clone RP11-2P7, WORKING DRAFT SEQUENCE, 38 unordered
            pieces.
ACCESSION   AC019261
VERSION     AC019261.3  GI:7229998
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 158405)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-2P7
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 158405)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C. Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 12, 2000 this sequence version replaced gi:6899774.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ------------------ Genome Center
               Center: Whitehead Institute/ MIT Center for Genome Research
               Center code: WIBR
               Web site: http://www-seq.wi.mit.edu
               Contact: sequence_submissions@genome.wi.mit.edu
            ------------------ Project Information
               Center project name: L2694
               Center clone name: 2_P_7
            ------------------ Summary Statistics
               Sequencing vector: M13; M77815; 100% of reads
               Chemistry: Dye-terminator Big Dye; 100% of reads
               Assembly program: Phrap; version 0.960731
               Consensus quality: 135380 bases at least Q40
               Consensus quality: 146343 bases at least Q30
               Consensus quality: 150850 bases at least Q20
               Insert size: 173000; agarose-fp
               Insert size: 154705; sum-of-contigs
               Quality coverage: 2.8 in Q20 bases; agarose-fp
               Quality coverage: 3.2 in Q20 bases; sum-of-contigs
            ------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 38 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
```

```
*  as soon as it is available and the accession number will
*  be preserved.
*         1    1607: contig of 1607 bp in length
*      1608    1707: gap of 100 bp
*      1708    3028: contig of 1321 bp in length
*      3029    3128: gap of 100 bp
*      3129    4889: contig of 1761 bp in length
*      4890    4989: gap of 100 bp
*      4990    6433: contig of 1444 bp in length
*      6434    6533: gap of 100 bp
*      6534    8085: contig of 1552 bp in length
*      8086    8185: gap of 100 bp
*      8186    9968: contig of 1783 bp in length
*      9969   10068: gap of 100 bp
*     10069   11857: contig of 1789 bp in length
*     11858   11957: gap of 100 bp
*     11958   14371: contig of 2414 bp in length
*     14372   14471: gap of 100 bp
*     14472   16047: contig of 1576 bp in length
*     16048   16147: gap of 100 bp
*     16148   18676: contig of 2529 bp in length
*     18677   18776: gap of 100 bp
*     18777   21785: contig of 3009 bp in length
*     21786   21885: gap of 100 bp
*     21886   24973: contig of 3088 bp in length
*     24974   25073: gap of 100 bp
*     25074   27441: contig of 2368 bp in length
*     27442   27541: gap of 100 bp
*     27542   30718: contig of 3177 bp in length
*     30719   30818: gap of 100 bp
*     30819   33992: contig of 3174 bp in length
*     33993   34092: gap of 100 bp
*     34093   36731: contig of 2639 bp in length
*     36732   36831: gap of 100 bp
*     36832   39283: contig of 2452 bp in length
*     39284   39383: gap of 100 bp
*     39384   41838: contig of 2455 bp in length
*     41839   41938: gap of 100 bp
*     41939   44209: contig of 2271 bp in length
*     44210   44309: gap of 100 bp
*     44310   47729: contig of 3420 bp in length
*     47730   47829: gap of 100 bp
*     47830   50393: contig of 2564 bp in length
*     50394   50493: gap of 100 bp
*     50494   56065: contig of 5572 bp in length
*     56066   56165: gap of 100 bp
*     56166   61306: contig of 5141 bp in length
*     61307   61406: gap of 100 bp
*     61407   65295: contig of 3889 bp in length
*     65296   65395: gap of 100 bp
*     65396   71547: contig of 6152 bp in length
*     71548   71647: gap of 100 bp
*     71648   77172: contig of 5525 bp in length
*     77173   77272: gap of 100 bp
*     77273   80143: contig of 2871 bp in length
*     80144   80243: gap of 100 bp
*     80243   84307: contig of 4064 bp in length
*     84308   84407: gap of 100 bp
*     84407   88430: contig of 4023 bp in length
*     88431   88530: gap of 100 bp
*     88531   94075: contig of 5545 bp in length
*     94076   94175: gap of 100 bp
*     94176  100299: contig of 6124 bp in length
*    100300  100399: gap of 100 bp
*    100400  107046: contig of 6647 bp in length
*    107047  107146: gap of 100 bp
*    107147  114514: contig of 7368 bp in length
*    114515  114614: gap of 100 bp
*    114615  122515: contig of 7901 bp in length
*    122516  122615: gap of 100 bp
*    122616  131415: contig of 8800 bp in length
*    131416  131515: gap of 100 bp
*    131516  138482: contig of 6967 bp in length
*    138483  138582: gap of 100 bp
*    138583  144615: contig of 6033 bp in length
*    144616  144715: gap of 100 bp
*    144716  158405: contig of 13690 bp in length.
```

```
FEATURES         Location/Qualifiers
     source          1..158405
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-2P7"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..1607
                     /note="assembly_fragment"
     misc_feature    1708..3028
                     /note="assembly_fragment"
     misc_feature    3129..4889
                     /note="assembly_fragment"
     misc_feature    4990..6433
                     /note="assembly_fragment"
     misc_feature    6534..8085
                     /note="assembly_fragment"
     misc_feature    8186..9968
                     /note="assembly_fragment"
     misc_feature    10069..11857
                     /note="assembly_fragment"
     misc_feature    11958..14371
                     /note="assembly_fragment"
     misc_feature    14472..16047
                     /note="assembly_fragment"
     misc_feature    16148..18676
                     /note="assembly_fragment"
     misc_feature    18777..21785
                     /note="assembly_fragment"
     misc_feature    21886..24973
                     /note="assembly_fragment"
     misc_feature    25074..27441
                     /note="assembly_fragment"
     misc_feature    27542..30718
                     /note="assembly_fragment"
     misc_feature    30819..33992
                     /note="assembly_fragment"
     misc_feature    34093..36731
                     /note="assembly_fragment"
     misc_feature    36832..39283
                     /note="assembly_fragment"
     misc_feature    39384..41838
                     /note="assembly_fragment"
     misc_feature    41939..44209
                     /note="assembly_fragment"
     misc_feature    44310..47729
                     clone_end:SP6
                     vector_side:left"
                     /note="assembly_fragment"
     misc_feature    47830..50393
                     /note="assembly_fragment"
     misc_feature    50494..56065
                     /note="assembly_fragment"
     misc_feature    56166..61306
                     /note="assembly_fragment"
     misc_feature    61407..65295
                     /note="assembly_fragment"
     misc_feature    65396..71547
                     /note="assembly_fragment
```

```
Query Match         27.0%;  Score 232;  DB 2;  Length 158405;
Best Local Similarity   99.6%;  Pred. No. 5.2e-114;
Matches 282;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy     23 CGGCCTTCGGCCGGGCCACCCAGCCGTGGTGCGGGGCGCTACCCGAGTCGGCTCTGCCAGC 82
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||| |||||||
Db 142537 CGGCCTTCGGCCGGGCCACCCAGCCGTGGTGCGGGGCGCTACCCGAGTCGGCTCGGCCAGC 142596

Qy     83 ACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTCGCCCGCGCGGAACGGCAGCACC 142
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     142597  ACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGCTCGCCCGCGCGGAACGGCAGCACC 142656
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     143     AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGGGCCTGCAGGTGGTGGAGCTGCCGGCCG 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142657  AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGGGCCTGCAGGTGGTGGAGCTGCCGGCCG 142716
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     203     ACGAGAGCCTTCCGACTGCGTCTTCGTGGAGGACGTGCCGTGGTGTGCGAGGAGACGG 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142717  ACGAGAGCCTTCCGACTGCGTCTTCGTGGAGGACGTGCCGTGGTGTGCGAGGAGACGG 142776
               |||||||||||||||||||||||||||||||||||||||||
Qy     263     CCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAGGAGGT 305
               ||||||||||||||||||||||||||||||||||||||
Db     142777  CCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAGGAGGT 142819
```

RESULT 12
AL360219/c
LOCUS   AL360219   168953 bp   DNA   linear   PRI 27-FEB-2001
DEFINITION Human DNA sequence from clone RP11-131L23 on chromosome 1, complete
            sequence.
ACCESSION   AL360219 AC073061
VERSION   AL360219.18 GI:13169534
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 168953)
  AUTHORS   Donnelly,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (27-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT   On or before May 15, 2001 this sequence version replaced
            gi:9958163, gi:12718084.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-131L23 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-131L23 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true left end of clone RP11-131L23 is at 168953 in this
            sequence. The true left end of clone RP4-611J7 is at 37497 in this
            sequence. The true right end of clone RP11-290M5 is at 104198 in
            this sequence. The true right end of clone RP4-621F18 is at 100 in
            this sequence. The true right end of clone RP4-611J7 is at 140647
            in this sequence.
FEATURES   Location/Qualifiers
     source   1..168953
                     /organism="Homo sapiens"
```

```
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-131L23"
                     /clone_lib="RPCI-11.1"
     misc_feature    13830..13855
                     /note="Sequence confirmed by AC073061 sequenced by WUGSC"
     misc_feature    27601..27792
                     /note="Sequence confirmed by AC019261 sequenced by WIBR"
     misc_feature    147415..147467
                     /note="Sequence from overlapping clone bA290M5 (AL162256).
                     Assembly confirmed by restriction digest."
     misc_feature    161180..161206
                     /note="Sequence from overlapping clone bA290M5 (AL162256).
                     Assembly confirmed by restriction digest."
ORIGIN
Query Match             27.0%;  Score 232;  DB 9;  Length 168953;
  Best Local Similarity   99.6%;  Pred. No. 5.2e-114;
  Matches  282;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy     23      CCGCCTTCGGCCGGCCACCCACGCCGTGGTGCGGGGCTACCCGAGTCGCTTCTGCCAGC 82
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74904   CGGCCTTCGGCCGGCCACCCACGCCGTGGTGCGGGGCTACCCGAGTCGCTTCTGCCAGC 74845
Qy     83      ACGGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTGCGCCCGCGCGGAACGGCAGCACC 142
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74844   ACGGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTGCGCCCGCGCGGAACGGCAGCACC 74785
Qy     143     AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGGGCCTGCAGGTGGTGAGCTGCCGGCCG 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74784   AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGGGCCTGCAGGTGGTGAGCTGCCGGCCG 74725
Qy     203     ACGAGAGCCTTCCGACTGCGTCTTCGTGGAGGACGTGCCGTGGTGTGCGAGGAGACGG 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74724   ACGAGAGCCTTCCGACTGCGTCTTCGTGGAGGACGTGCCGTGGTGTGCGAGGAGACGG 74665
Qy     263     CCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAGGAGGT 305
               ||||||||||||||||||||||||||||||||||||||||
Db     74664   CCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAGGAGGT 74622
```

RESULT 13
HSA123M24/c
LOCUS   HSA123M24   83946 bp   DNA   linear   PRI 04-MAR-2003
DEFINITION Human DNA sequence from clone RP4-621F18 on chromosome 1p11.4-21.3,
            complete sequence.
ACCESSION   AL078459
VERSION   AL078459.8 GI:5791502
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 83946)
  AUTHORS   Whitehead,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequest@sanger.ac.uk
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT   On Aug 27, 1999 this sequence version replaced gi:5596949.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep ------------
            Genome Center

Best Local Similarity 100.0%; Pred. No. 3.8e-34;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    302  AGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAG  361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    336  AGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAG  395

Qy    362  ATGAAAATGCAACTTTAGATGGCGG  386
           |||||||||||||||||||||||||
Db    396  ATGAAAATGCAACTTTAGATGGCGG  420
```

RESULT 15
BD034010
LOCUS   BD034010   420 bp   DNA   linear   PAT 27-AUG-2002
DEFINITION   Sequence tag and encoded human protein.
ACCESSION   BD034010
VERSION   BD034010.1 GI:22575752
KEYWORDS   JP 2001269182-A/10256.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 420)
AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE   Sequence tag and encoded human protein
JOURNAL   Patent: JP 2001269182-A 10256 02-OCT-2001;
             GENSET
COMMENT   OS   Homo sapiens (human)
             PN   JP 2001269182-A/10256
             PD   02-OCT-2001
             PF   24-FEB-2000 JP 2000118773
             PR   26-FEB-1999 US   60/122487
             PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
             PI   JORDAN
             PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
             PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
             G06F15/40
             CC
             FH   Key   Location/Qualifiers
FEATURES   Location/Qualifiers
     source   1..420
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   9.9%; Score 85; DB 6; Length 420;
Best Local Similarity 100.0%; Pred. No. 3.8e-34;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    302  AGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAG  361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    336  AGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAG  395

Qy    362  ATGAAAATGCAACTTTAGATGGCGG  386
           |||||||||||||||||||||||||
Db    396  ATGAAAATGCAACTTTAGATGGCGG  420
```

Search completed: June 7, 2004, 22:28:48
Job time : 3414 secs

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the rare
occasion of the clone being a YAC.
This sequence was generated from part of bacterial clone contigs of
human chromosome 1, constructed by the Sanger Centre Chromosome 1
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP4-621F18 is from the library RPCI-4 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.
FEATURES   Location/Qualifiers
     source   1..83946
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p11.4-21.3"
                     /clone="RP4-621F18"
                     /clone_lib="RPCI-4"
ORIGIN

Query Match   17.2%; Score 148; DB 9; Length 83946;
Best Local Similarity 100.0%; Pred. No. 2e-68;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     596  AGATCATGCAACAGATGAGTGACCACCGTACGACAAACTCACTGTGCCTGATGACATAG  655
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   18612  AGATCATGCAACAGATGAGTGACCACCGTACGACAAACTCACTGTGCCTGATGACATAG  18553

Qy     656  CAGCAAACTGTATATATCTAAATATCCCCAACAAAGGGCACGTCTTGCTGCACCGAACCC  715
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   18552  CAGCAAACTGTATATATCTAAATATCCCCAACAAAGGGCACGTCTTGCTGCACCGAACCC  18493

Qy     716  CGGAAGAGTATCCAGAAAGTGCAAAGGT  743
            ||||||||||||||||||||||||||||
Db   18492  CGGAAGAGTATCCAGAAAGTGCAAAGGT  18465
```

RESULT 14
AX898477
LOCUS   AX898477   420 bp   DNA   linear   PAT 18-DEC-2003
DEFINITION   Sequence 14340 from Patent EP1033401.
ACCESSION   AX898477
VERSION   AX898477.1 GI:40053390
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE   Expressed sequence tags and encoded human proteins
JOURNAL   Patent: EP 1033401-A 14340 06-SEP-2000;
             Genset (FR)
FEATURES   Location/Qualifiers
     source   1..420
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   9.9%; Score 85; DB 6; Length 420;

```
                  GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on:     June  7, 2004, 15:45:50 ; Search time 405 Seconds
                              (without alignments)
                              8999.886 Million cell updates/sec

Title:         US-09-889-733B-1
Perfect score: 858
Sequence:      1 atggccggcctcggccaccc........acaagaaggtagactcctga 858

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:    6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    N_Geneseq_29Jan04:*
              1: geneseqn1980s:*
              2: geneseqn1990s:*
              3: geneseqn2000s:*
              4: geneseqn2001as:*
              5: geneseqn2001bs:*
              6: geneseqn2002s:*
              7: geneseqn2003as:*
              8: geneseqn2003bs:*
              9: geneseqn2003cs:*
             10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

```
Result         %
No.    Score   Query Match  Length DB ID        ID        Description
                                                          
 1      858   100.0    858     3   AAA47655    AAA47655   Aaa47655 Dimethyla
 2     851.6   99.3    858     9   ADE71465    ADE71465   Ade71465 Human Thr
 3     851.6   99.3   1633     4   AAH02907    AAH02907   Aah02907 Human she
 4     851.6   99.3   1633     6   ABQ88102    ABQ88102   Abq88102 Human ost
 5     851.6   99.3   3932     7   ABZ35939    ABZ35939   Abz35939 Human sec
 6     851.6   99.3   4208     6   AAL41494    AAL41494   Aal41494 Drug meta
 7      850    99.1    858     6   ADE71463    ADE71463   Ade71463 Human Met
 8     715.6   83.4   3008     9   ADB53014    ADB53014   Adb53014 Primary r
 9     715.6   83.4   3008     9   ADB79843    ADB79843   Adb79843 Rat N-G,N
10     705.4   82.2   3778     6   ADE53981    ADE53981   Ade53981 Human pro
11     630.4   73.5   3651     3   AAF16077    AAF16077   Aaf16077 Human pro
12     218.8   25.5    444     6   ABN89945    ABN89945   Abn89945 Mouse clo
13     216.2   25.2   1376     2   AAZ52967    AAZ52967   Aaz52967 Human pro
14     214.6   25.0    858     3   AAA47656    AAA47656   Aaa47656 Dimethyla
15     214.6   25.0   1228     3   AAD00677    AAD00677   Aad00677 Human Hyd
16     207.6   24.2    845     3   AAA99390    AAA99390   Aaa99390 Human mel
17     202.6   23.6    862     6   ABQ47888    ABQ47888   Abq47888 Oligonucl
18     202.6   23.6    862     6   ABQ47889    ABQ47889   Abq47889 Oligonucl
19     202.6   23.6    898     6   ABQ41319    ABQ41319   Abq41319 Oligonucl
20     202.6   23.6    898     6   ABQ41318    ABQ41318   Abq41318 Oligonucl
21     197.8   23.1    770     6   ABS76755    ABS76755   Abs76755 Frog embr
22     167.8   19.6    862     6   ABQ47886    ABQ47886   Abq47886 Oligonucl
23     167.8   19.6    862     6   ABQ47887    ABQ47887   Abq47887 Oligonucl
```

```
24          167.8   19.6    898     6   ABQ41321   ABQ41321  Oligonucl
25          167.8   19.6    898     6   ABQ41320   ABQ41320  Oligonucl
26          148.4   17.3    472     8   ACH21905   ACH21905  Human adu
27   c      148     17.2   83946    6   ABQ88101   ABQ88101  Human ost
28          146     17.0    452     8   ACH27659   ACH27659  Human adu
29          134.2   15.6    403     8   ACH32042   ACH32042  Human end
30          117.4   13.7   6373     5   ABA14642   ABA14642  Human ner
31          117.4   13.7   6381     5   ABA14641   ABA14641  Human ner
32          107.4   12.5    873     4   ABL03335   ABL03335  Drosophil
33   c      103.4   12.1    655     9   ADD34806   ADD34806  Mouse mit
34           85              420    3   AAC10265   AAC10265  Human sec
35   c       81.2    9.5    239     5   AAF68469   AAF68469  Human lun
36   c       81.2    9.5    239     7   ABK38380   ABK38380  cDNA enco
37   c       81.2    9.5    239     7   ACA10709   ACA10709  Human lun
38   c       81.2    9.5    239     7   ABX99660   ABX99660  Lung canc
39   c       81.2    9.5    239    10   ADE71675   ADE71675  Human lun
40           66.2    7.7   3643     4   ABL03334   ABL03334  Drosophil
41           65.8    7.7   3223     4   AAK73698   AAK73698  Human imm
42           65.8    7.7  28772     9   ABK83555   ABK83555  Human cDN
43   c       65.2    7.6    504     9   ADB57724   ADB57724  Toxicity-
44   c       63.2    7.4    545     9   ADD34805   ADD34805  Mouse mit
45           60.8    7.1   2000     7   ADA71938   ADA71938  Rice gene
```

ALIGNMENTS

```
RESULT 1
AAA47655
ID   AAA47655 standard; cDNA; 858 BP.
XX
AC   AAA47655;
XX
DT   08-NOV-2000  (first entry)
XX
DE   Dimethylarginine dimethylaminohydrolase (DDAH1) coding sequence.
XX
KW   Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH2;
KW   arginine deaminase; hyperlipidemia; renal failure; hypertension;
KW   restenosis; atherosclerosis; multiple sclerosis; cancer;
KW   ischemia reperfusion injury; septic shock; multi organ failure;
KW   arthritis; skin disorders; inflammatory cardiac disease; migraine;
KW   infection; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..858
FT                   /*tag= a
FT                   /product= "DDAH1"
XX
PN   WO20004888-A2.
XX
PD   03-AUG-2000.
XX
DR   26-JAN-2000; 2000WO-GB000226.
XX
PF   
PR   26-JAN-1999;     99GB-00001705.
PR   04-JUN-1999;     99GB-00013066.
XX
PA   (UNLO ) UNIV COLLEGE LONDON.
XX
PI   Vallance PJT, Leiper JM, Whitley GSJ, Charles IG;
XX
DR   WPI; 2000-543392/49.
DR   P-PSDB; AAB01477.
XX
PT   Novel methylarginase polypeptides and polynucleotides, used to identify
PT   modulators of them, which are used in the treatment of e.g. cancer,
PT   hypertension, and bacterial infections.
XX
PS   Claim 1; Page 55-56; 68pp; English.
XX
```

CC   Nucleotides encoding methylarginase polypeptides, vectors comprising
CC   these nucleotides and the polypeptides themselves can be used in
CC   medicaments for the treatment of hyperlipidemia, renal failure,
CC   hypertension, restenosis after angioplasty, atherosclerosis,
CC   complications of heart failure, schizophrenia, multiple sclerosis or
CC   cancer. Modulators of the enzyme can be used in medicaments for the
CC   treatment of ischemia-reperfusion injury of the brain or heart, cancer,
CC   lethal hypertension in severe inflammatory conditions such as septic
CC   shock or multi-organ failure, or local and systemic inflammatory
CC   disorders including arthritis, skin disorders, inflammatory cardiac
CC   disease, migraine, or microbial or bacterial infection
XX
SQ   Sequence 858 BP; 217 A; 214 C; 256 G; 171 T; 0 U; 0 Other;

Query Match   100.0%; Score 858; DB 3; Length 858;
Best Local Similarity   100.0%; Pred. No. 1.5e-185;
Matches 858; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 ATGGCCGGCTCGGCCACCCTCGCCTTCGGCGGGCCACCCACGCCGTGGTGCGGCGGCG   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGGCCGGCTTCGGCCACCCTCGCCTTCGGCGGGCCACCCACGCCGTGGTGCGGCGGCG   60

Qy   61 CTACCCGAGTGCGTCTGCCAGCACGCGCTGAGAAGGGCAGCAAGGGCGAGGAGGTGGACGTC  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTACCCGAGTGCGTCTGCCAGCACGCGCTGAGAAGGGCAGCAAGGGCGAGGAGGTGGACGTC  120

Qy  121 GCCCGCCGCGGAACGGCAGCGACCAGCTCACGTGGGGCTGTGGGCAGCAAGCTGGGGCTG  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GCCCGCCGCGGAACGGCAGCGACCAGCTCACGTGGGGCTGTGGGCAGCAAGCTGGGGGCTG  180

Qy  181 CAGTGGTGGAGTGCCGGCCGACGAGAGCCTTCCGGACTGCCTCTTCGTGAGGACGTG  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CAGTGGTGGAGTGCCGGCCGACGAGAGCCTTCCGGACTGCCTCTTCGTGAGGACGTG  240

Qy  241 GCCGGTGGTGTGGAGGAGACGGGCCCTCATCACCCGGCGCCGGAGCCCGAGCCGGAGGAAG  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GCCGGTGGTGTGGAGGAGACGGGCCCTCATCACCCGGCGCCGGAGCCCGAGCCGGAGGAAG  300

Qy  301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA  360

Qy  361 GATGAAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTTGTG  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GATGAAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTTGTG  420

Qy  421 GGCCTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GGCCTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC  480

Qy  481 TATGCAGTCTCCACAGTGCCAGTGGCAGTGGGTTGCATTTGAAGAGTTTCTGCAGCATG  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TATGCAGTCTCCACAGTGCCAGTGGCAGTGGGTTGCATTTGAAGAGTTTCTGCAGCATG  540

Qy  541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCCTTAAGATC  600

Qy  601 ATGCAACAGATGAGTGACCACCGCTACAGACAAACTCACTGTGCCTGATGACATAGCAGCA  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ATGCAACAGATGAGTGACCACCGCTACAGACAAACTCACTGTGCCTGATGACATAGCAGCA  660

Qy  661 AACTGTATATCTAAATATCCCAACAAAAGGGCACGTCTTGCTGCCGACCCCGGAA  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AACTGTATATCTAAATATCCCAACAAAAGGGCACGTCTTGCTGCCGACCCCGGAA  720

Qy  721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCCGTG  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCCGTG  780

Qy  781 AGCATGTCTGAACTGGAAAAGGTGGATGGCCTGCTCACCTGTGCTCAGTTTTAATTAAC  840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AGCATGTCTGAACTGGAAAAGGTGGATGGCCTGCTCACCTGTGCTCAGTTTTAATTAAC  840

Qy  841 AAGAAGGTAGACTCCTGA  858
        ||||||||||||||||||
Db  841 AAGAAGGTAGACTCCTGA  858
```

RESULT 2
ADE71465
ID   ADE71465 standard; DNA; 858 BP.
XX
AC   ADE71465;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human Thr87 wild-type dimethylarginine dimethylaminohydrolase-1 DNA.
XX
KW   Human; dimethylarginine dimethylaminohydrolase-1; DDAH 1; chromosome 1;
KW   single nucleotide polymorphism; SNP; ADMA;
KW   asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
KW   coronary heart disease; cerebrovascular disease; hypertension; diabetes;
KW   susceptibility; genotyping; therapeutic targeting; antidiabetic;
KW   cardiant; cerebroprotective; hypotensive; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..858
FT   /tag= a
FT   /product= "wild-type (Thr87) DDAH 1"
FT   variation   replace(260,T)
FT   /tag= b
FT   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   /note= "Results in Thr (C variant) or Met (T variant) at
FT   position 87 of the protein"
XX
PN   WO200308963889638-A1.
XX
PD   30-OCT-2003.
XX
PF   11-APR-2003; 2003WO-FI000274.
XX
PR   19-APR-2002; 2002US-00155456.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Valkonen V, Salonen JT, Pirskanen M, Tuomainen T, Laakso J;
PI   Laaksonen R;
XX
DR   WPI; 2003-854121/79.
DR   P-PSDB; ADE71466.
XX
PT   New nucleic acid encoding a variant dimethylarginine
PT   dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
PT   and its vascular complications, e.g. coronary or cerebrovascular disease
PT   or hypertension.
XX
PS   Disclosure; SEQ ID NO 3; 37pp; English.
XX
CC   The invention relates to a variant dimethylarginine
CC   dimethylaminohydrolase-1 (DDAH 1; ADE71464) and a nucleic acid encoding
CC   it (ADE71463). DDAHs regulate the metabolism of ADMA (asymmetrical
CC   dimethyl-arginine; a naturally occurring nitric oxide synthase inhibitor)
CC   and nitric oxide, which are relevant in disorders such as cardiovascular
CC   disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
CC   has been implicated in susceptibility to diseases such as familial
CC   combined hyperlipidaemia, premature coronary artery disease (CAD), non
CC   insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
CC   The variant DDAH 1 of the invention has Met at position 87, rather than
CC   Thr, which is caused by a C to T change in exon 1 at a position
CC   corresponding to base 260 in ADE71463 and ADE71465. Individuals hetero-
CC   or homozygous for the Thr87Met allele have an increased risk of
CC   developing cardiovascular diseases and diabetes. The invention also
CC   relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a

```
CC   susceptibility to a cardiovascular disease and diabetes in an individual
CC   by determining the DDAH 1 genotype; a method of treating diabetes or
CC   vascular complications of diabetes by enhancing nitric oxide
CC   availability, production or concentration; a method of targeting
CC   treatment of cardiovascular disease and diabetes in a hypertensive
CC   patient by determining the DDAH 1 genotype and treating them with a drug
CC   that affects nitric oxide availability, production or metabolism; and a
CC   transgenic animal which carries a human variant DDAH 1 nucleic acid
CC   sequence. The nucleic acid molecules and polypeptides are useful for
CC   treating cardiovascular disease (e.g., coronary heart disease,
CC   cerebrovascular disease, and hypertension), and diabetes and its vascular
CC   complications. The methods are useful for determining whether a patient
CC   will benefit from treatment with a drug which affects nitric oxide
CC   availability, production or metabolism; a drug which reduces ADMA
CC   availability or concentration; orn an agent which elevates DDAH
CC   availability or concentration (such as DDAH agonist). The methods are
CC   also useful for determining whether a patient will be at risk of adverse
CC   effects if DDAH antagonists are administered. The present sequence
CC   represents a nucleic acid encoding the "wild-type" (Thr87) DDAH 1.
XX
SQ   Sequence 858 BP; 218 A; 213 C; 258 G; 169 T; 0 U; 0 Other;

Query Match             99.3%;  Score 851.6;  DB 9;  Length 858;
Best Local Similarity   99.5%;  Pred. No. 4.3e-184;
Matches  854;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

Qy          1 ATGGCCGGCCTCGGCCACCCCTCCGCCTTCGGCCGGCGGCCACCCACGCCGTGGTGCGGGCG 60
              ||||||||| ||||| ||||| ||||||||||||||| ||||||||||||||||||||||||
Db          1 ATGGCCGGGCTCGGCCACCCCGCCGCCTTCGGCCGGCCACCCACGCCGTGGTGCGGGCG 60

Qy         61 CTACCCGAGTCGCTCTGCCAGCACGCGCTGGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTACCCGAGTCGCTCGGCCAGCACGCGCTGGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 120

Qy        121 GCCCGCGGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGTCGGGGCTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCCGCGGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGTCGGGGCTG 180

Qy        181 CAGGTGGTGAGCTGCCGGCCGACGAGAGCCTTCGGACGTCTTCGTGGAGGACGTG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGTGGTGAGCTGCCGGCCGACGAGAGCCTTCGGACGTCTTCGTGGAGGACGTG 240

Qy        241 GCCTGGTGTGCGAGGAGACGGGCCCTCATCACCCGACCCGGGCGCGGAGAAG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGTGGTGTGCGAGGAGACGGGCCCTCATCACCCGACCCGGGCGCCGAGGAAG 300

Qy        301 GAGGTTGACATGATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGGTTGACATGATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360

Qy        361 GATGAAAATGCAACTTAGATGGCCGAGAGTGTTTATTCACAGGCAGAGAATTTTTTGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGAAAATGCAACTTAGATGGCCGAGAGTGTTTATTCACAGGCAGAGAATTTTTTGTG 420

Qy        421 GGCCTTTCAAAAGGACAAATCACCGAGGTGCTGAAAATCTTGGCTGATACTTTTAAGGAC 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCCTTTCAAAAGGACAAATCACCGAGGTGCTGAAAATCTTGGCTGATACTTTTAAGGAC 480

Qy        481 TATGCAGTCTCCACAGTGCCAGTGCAGTGGGTTGCATTGAAGAGTTTCTGCAGCATG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TATGCAGTCTCCACAGTGCCAGTGCAGTGGGTTGCATTGAAGAGTTTCTGCAGCATG 540

Qy        541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600

Qy        601 ATGCAACAGATGAGTGACCACGTGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATGCAACAGATGAGTGACCACGTGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660

Qy        661 AACTGTATATATCTAAATATCCCACAAAAGGGCACGTCTTGCTCACCGAACCCGGAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AACTGTATATATCTAAATATCCCACAAAAGGGCACGTCTTGCTCACCGAACCCGGAA 720
```

```
Qy        721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG 780

Qy        781 AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC 840

Qy        841 AAGAAGGTAGACTCCTGA 858
              ||||| ||||||||||||
Db        841 AAGAAAGTAGACTCCTGA 858


RESULT 3
AAH02907
ID   AAH02907 standard; DNA; 1633 BP.
XX
AC   AAH02907;
XX
DT   15-JUN-2001 (first entry)
XX
DE   Human shear stress-response coding sequence SEQ ID NO: 67.
XX
KW   Human; shear stress-response protein; vascular disease; arteriosclerosis;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200125427-A1.
XX
PD   12-APR-2001.
XX
PF   02-OCT-2000; 2000WO-JP006840.
XX
PR   01-OCT-1999;   99JP-00280976.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
PA   (NOJI/) NOJIMA H.
XX
PI   Nojima H,  Yoshisue H,  Obayashi M,  Ota T,  Kawabata A,  Sakurada K;
PI   Kuga T,  Sekine S,  Nakamura Y,  Sugano S;
XX
DR   WPI; 2001-266308/27.
DR   P-PSDB; AAB90784.
XX
PT   DNA sequences, proteins encoded by them and antibodies against them
PT   useful in diagnosis and treatment of vascular disease caused by
PT   arteriosclerosis.
XX
PS   Claim 20; Page 404-407; 678pp; Japanese.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human shear stress response proteins. These are useful in the
CC   diagnosis, treatment and screening of vascular diseases caused by
CC   arteriosclerosis, including heart failure, post-PTCA restenosis and
CC   hypertension
XX
SQ   Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Query Match             99.3%;  Score 851.6;  DB 4;  Length 1633;
Best Local Similarity   99.5%;  Pred. No. 5.1e-184;
Matches  854;  Conservative   0;  Mismatches   4;  Indels   4;  Gaps   0;

Qy          1 ATGGCCGGCCTCGGCCACCCCTCCGCCTTCGGCCGGCGGCCACCCACGCCGTGGTGCGGGCG 60
              ||||||||| ||||| ||||| ||||||||||||||| ||||||||||||||||||||||||
Db        323 ATGGCCGGGCTCGGCCACCCCGCCGCCTTCGGCCGGCCACCCACGCCGTGGTGCGGGCG 382

Qy         61 CTACCCGAGTCGCTCTGCCAGCACGCGCTGGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 120
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTACCCGAGTCGCTCGGCCAGCACGCGCTGGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 442

Qy        121 GCCCGCGGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGTCGGGGCTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 CTACCCGAGTCGCTCGGCCAGCACGCGCTGGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 442

Qy        121 GCCCGCGGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGTCGGGGCTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCCGCGGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGTCGGGGCTG 180
```

PR   18-DEC-2000; 2000US-0255882P.
PR   24-APR-2001; 2001US-0285691P.
XX
PA   (GENE-) GENE LOGIC INC.
PA   (PROC) PROCTER & GAMBLE CO.
XX
PI   Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI   Mertz L;
XX
DR   WPI; 2002-557663/59.
XX
PT   Use of genes and their expression profiles associated with osteoblast
PT   differentiation for screening modulators bone formation, for diagnosing
PT   or treating e.g. osteoporosis, or as markers for the differentiation
PT   process.
XX
PS   Claim 1; SEQ ID NO 9; 78pp + Sequence Listing; English.
XX
CC   The invention relates to genes and their expression profiles are used
CC   for: (a) screening modulators of precursor stem cell differentiation into
CC   osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC   deposition of bone tissue, abnormal rate of osteoblast formation or
CC   osteoporosis; or (c) treating or monitoring treatment of the conditions
CC   cited in (b), or monitoring the progression of bone tissue deposition.
CC   Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC   osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC   induced abnormalities in bone formation or bone loss, conditions that
CC   involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC   skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC   or fibrous dysplasia. The present sequence is that of an osteoblast
CC   differentiation associated cDNA marker of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Query Match   99.3%; Score 851.6; DB 6; Length 1633;
Best Local Similarity 99.5%; Pred. No. 5.1e-184;
Matches 854; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1 ATGGCCGGCTCCGGCCACCCTCCGGCCTTCGGCCGGCCACCCAGCGCCGTGGTGCGGGCG 60
Db 323 ATGGCCGGGCTCGGGCCACCCCCGGCCTTCGGCCGGCCACCCACCCAGCCGTGGTGCGGGCG 382

Qy   61 CTACCCGAGTCGCTCTGCGCCAGCACGCGGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 120
Db 383 CTACCCGAGTCGCTCTGCGCCAGCACGCGGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 442

Qy 121 GCCGGCGCCGGAACGGCAGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 180
Db 443 GCCGGCGCCGGAACGGCAGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 502

Qy 181 CAGGTGGTGGAGCTGCCGGACTGCGTCTTCGTGGAGGACGTG 240
Db 503 CAGGTGGTGGAGCTGCCGGACTGCGTCTTCGTGGAGGACGTG 562

Qy 241 GCCGTGGTGTGCGAGGAGGACGGCCCTCATCACCCGAGCCGCGGAGGAAG 300
Db 563 GCCGTGGTGTGCGAGGAGGACGGCCCTCATCACCCGAGCCGCGGAGGAAG 622

Qy 301 GAGGTTGACATGATGAAGAAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
Db 623 GAGGTTGACATGATGAAGAAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 682

Qy 361 GATGAAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGACGAGAATTTTTTGTG 420
Db 683 GATGAAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGACGAGAATTTTTTGTG 742

Qy 421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
Db 743 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 802

DR   WPI; 2003-075543/07.
DR   P-PSDB; ABP75492.
XX
PT   New human secretory proteins and polynucleotides, useful for diagnosing,
PT   treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
PT   neurological disorders (e.g. Alzheimer's), or cell proliferations or
PT   cancers.
XX
PS   Claim 1; SEQ ID NO 103; 45bpp + Sequence Listing; English.
XX
CC   The invention relates to a secretory polynucleotide (designated sptm)
CC   comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
CC   naturally occurring polynucleotide sequence at least 90 % identical to
CC   the polynucleotide sequence, a polynucleotide complementary to them or an
CC   RNA equivalent of them. The polypeptide or polynucleotide are useful for
CC   treating, preventing or diagnosing a disease or condition associated with
CC   the expression of functional SPTM. These are particularly useful for
CC   diagnosing, treating or preventing autoimmune/inflammatory disorders
CC   (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
CC   disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC   dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
CC   multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
CC   schizophrenia or amnesia), or cell proliferative disorders (e.g.
CC   psoriasis, polycythemia vera, or cancers including adenocarcinoma,
CC   leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
CC   breast, cervix or prostate). Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3932 BP; 1089 A; 883 C; 775 G; 1185 T; 0 U; 0 Other;

Query Match   99.3%; Score 851.6; DB 7; Length 3932;
Best Local Similarity   99.5%; Pred. No. 6.5e-184;
Matches 854; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   1 ATGGCCGGCCTCGGCCACCCCTCCGCCTTCGGCCGCGGCCACCCACGCGGTGGTGCGGGCG 60
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ATGGCCGGGCTCGGCCACCCCTCCGCCTTCGGCCGCGGCCACCCACGCGGTGGTGCGGGCG 76

Qy   61 CTACCCGAGTCGCTGCTGCCAGCACGCGGTCGAAGAGGCGCCAAGGCGAGGAGGTGGACGTC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 CTACCCGAGTCGCTGCTGCCAGCACGCGGTCGAAGAGGCGCCAAGGCGAGGAGGTGGACGTC 136

Qy   121 GCCCGGCGGGAACGGCAGGCAGCACCAGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 GCCCGGCGGGAACGGCAGGCAGCACCAGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 196

Qy   181 CAGGTGGTGGAGTGCCGGCCGGACGAGACCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   197 CAGGTGGTGGAGTGCCGGCCGGACGAGACCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 256

Qy   241 GCCGTGGTGTGCGGAGGAGGACGGCCCTCATCACCCGACACCCGGGCGCCGGAGGAG 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   257 GCCGTGGTGTGCGGAGGAGGACGGCCCTCATCACCCGACACCCGGGCGCCGGAGGAG 316

Qy   301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGGATGATGAAA 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   317 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGGATGATGAAA 376

Qy   361 GATGAAAATGCAACTTGAGTGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   377 GATGAAAATGCAACTTGAGTGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 436

Qy   421 GGCCTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
             |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   437 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 496

Qy   481 TATGCAGTCTCCAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATG 540
             ||||||||||||| |||||||||||||||||||||| ||||||||||||||||||||
Db   497 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATG 556

Qy   541 GCTGGGCCTAACCTGATCCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600

Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   803 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 862

Qy   541 GCTGGGCCTAACCTGATCCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   863 GCTGGGCCTAACCTGATCCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 922

Qy   601 ATGCAACAGATGAGGTGACCACCGCTACGACGAAACTCACTGTGCCTGATGACATAGCAGCA 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   923 ATGCAACAGATGAGGTGACCACCGCTACGACGACAAACTCACTGTGCCTGATGACATAGCAGCA 982

Qy   661 AACTGTATATATCTAAATATCCCCAACAAAGGGCACGTCTTGCTGCACCGACCCCCGGAA 720
             ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   983 AACTGTATATATCTAAATATCCCCAACAAAGGGCACGTCTTGCTGCACCGACCCCGGAA 1042

Qy   721 GAGTATCCAGAAAGTGCAAAGGTTTATGAAGAAACTGAAGGACCATATGCTGATCCCCGTG 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1043 GAGTATCCAGAAAGTGCAAAGGTTTATGAAGAAACTGAAGGACCATATGCTGATCCCCGTG 1102

Qy   781 AGCATGTCTGAACTGGAAAAGGTGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1103 AGCATGTCTGAACTGGAAAAGGTGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC 1162

Qy   841 AAGAAGTAGACTCCTGA 858
             |||||||||||||||||
Db   1163 AAGAAGTAGACTCCTGA 1180

RESULT 5
AB235939
ID   AB235939 standard; cDNA; 3932 BP.
XX
AC   AB235939;
XX
DT   10-FEB-2003 (first entry)
XX
DE   Human secretory polynucleotide SPTM SEQ ID NO 103.
XX
KW   Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
KW   asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
KW   Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
KW   multiple sclerosis; Parkinson's disease; cell proliferative disorder;
KW   anti-inflammatory; immunosuppressive; neuroprotective; nootropic;
KW   neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
KW   antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
KW   secretory polynucleotide; secretory protein; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200283876-A2.
XX
PD   24-OCT-2002.
XX
PF   27-MAR-2002; 2002WO-US009921.
XX
PR   29-MAR-2001; 2001US-0280067P.
PR   29-MAR-2001; 2001US-0280068P.
PR   16-MAY-2001; 2001US-0291280P.
PR   17-MAY-2001; 2001US-0291829P.
PR   17-MAY-2001; 2001US-0291849P.
PR   19-JUN-2001; 2001US-0299428P.
PR   20-JUN-2001; 2001US-0299776P.
PR   20-JUN-2001; 2001US-0300001P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI   Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI   Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI   Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI   Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX

```
Db  557 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTAAGATC 616
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  601 ATGCCAACAGATGAGTGACCACCCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  617 ATGCCAACAGATGAGTGACCACCCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 676

Qy  661 AACTGTATATATCTAAATATCCCAACAAAAGGGCACGTCTTGCTCGCACCGAACCCCGGAA 720
            |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  677 AACTGTATATATCTAAATATCCCAACAAAAGGGCACGTCTCTGCGCACCGAACCCCGGAA 736

Qy  721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  737 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG 796

Qy  781 AGCATGTCTGAACTGGAAAAAGTGGATGGGCTGCTCACCCGCTGCTGCTCAGTTTAATTAAC 840
            ||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db  797 AGCATGTCTGAACTGGAAAAAGTGGATGGGCTGCTCACCCTGCTGCTCAGTTTTAATTAAC 856

Qy  841 AAGAAGGTAGACTCCTGA 858
            ||||| ||||||||||
Db  857 AAGAAGTAGACTCCTGA 874
```

RESULT 6
AAL41494
ID AAL41494 standard; DNA; 4208 BP.
XX
AC AAL41494;
XX
DT 21-NOV-2002 (first entry)
XX
DE Drug metabolising enzyme encoding DNA - 7671089CB1.
XX
KW Anti-HIV; antiarteriosclerotic; dermatological; cytostatic; thyromimetic;
KW osteopathic; thrombolytic; ophthalmological; antilipaemic; hepatotropic;
KW antidiarrhoeic; antiinflammatory; virucide; immunogenic; autoimmune;
KW contact dermatitis; cell-proliferative; cancer; cirrhosis; dwarfism;
KW developmental; hypothyroidism; endocrine; osteoporosis; thrombosis;
KW diabetes; glaucoma; keratitis; metabolic; hyperlipidaemia; diarrhoea;
KW cystic fibrosis; gastrointestinal; gastroenteritis; liver; hepatitis;
KW Reye's syndrome; exogenous compound; gene therapy; enzyme; human; gene;
KW ds.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   200..1225
FT   /tag= a
FT   /product= "Drug metabolising enzyme protein"
XX
PN WO20026654-A2.
XX
PD 29-AUG-2002.
XX
PF 14-FEB-2002; 2002WO-US004918.
XX
PR 16-FEB-2001; 2001US-0269643P.
PR 23-FEB-2001; 2001US-0271332P.
PR 16-MAR-2001; 2001US-0276767P.
PR 06-APR-2001; 2001US-0282077P.
PR 19-APR-2001; 2001US-0285447P.
PR 27-APR-2001; 2001US-0287060P.
PR 03-MAY-2001; 2001US-0288543P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Astromoff A, Au-Young J, Baughn MR, Ding L, Duggan BM;
PI Forsythe IJ, Gietzen KJ, Griffin JA, Lee EA, Lu Y, Richardson TW;
PI Ring HZ, Sanjanwala MM, Swarnakar A, Walia NK, Warren BA, Xu Y;
PI Yue H, Zebarjadian Y;
XX
DR WPI; 2002-674949/72.

DR P-PSDB; AAO22798.
XX
PT New drug metabolizing enzymes (DME) useful for diagnosing, treating and
PT preventing diseases or conditions associated with aberrant DME
PT expression, e.g. cancer, AIDS, atherosclerosis, diabetes, glaucoma,
PT hepatitis, osteoporosis.
XX
PS Claim 5; Page 164-165; 166pp; English.
XX
CC The invention relates to an isolated polypeptide comprising 12 fully
CC defined sequences of 81-615 amino acids given in the specification; a
CC naturally occurring amino acid sequence at least 90% identical to, having
CC 81-599 amino acids, at least 93% identical to a sequence of 529 amino
CC acids, or at least 97% identical to a sequence of 615 amino acids, all
CC given in the specification; or a biologically active or immunogenic
CC fragment of the polypeptide. The polypeptides and polynucleotides are
CC useful in diagnosing, treating and preventing diseases or conditions
CC associated with the decreased expression or overexpression of a drug
CC metabolising enzyme (DME), such as autoimmune/inflammatory (e.g. AIDS,
CC atherosclerosis, contact dermatitis) cell-proliferative (e.g. cancer,
CC cirrhosis), developmental (e.g. dwarfism, hypothyroidism), endocrine
CC (e.g. osteoporosis, thrombosis, diabetes), eye (e.g. glaucoma,
CC keratitis), metabolic (e.g. hyperlipidaemia, cystic fibrosis)
CC gastrointestinal (e.g. gastroenteritis, diarrhoea), or liver (e.g.
CC hepatitis, Reye's syndrome) disorders. These are also useful in assessing
CC the effects of exogenous compounds on the expression of nucleic acid and
CC amino acid sequences of DME. The DME or its fragments are useful in
CC screening compounds for effectiveness as agonist or antagonist of the
CC polypeptides, or in altering the expression of the target polynucleotide
CC and compounds that specifically bind to or modulate the activity of the
CC polypeptide. The microarray is useful in monitoring or measuring protein-
CC protein interactions, drug-target interactions, and gene expression
CC profiles. The polynucleotides of the invention can be used in gene
CC therapy. This polynucleotide sequence represents the DNA encoding a drug
CC metabolising enzyme of the invention
XX
SQ Sequence 4208 BP; 1092 A; 1017 C; 883 G; 1216 T; 0 U; 0 Other;

Query Match   99.3%;   Score 851.6; DB 6;   Length 4208;
Best Local Similarity   99.5%;   Pred. No. 6.6e-184;
Matches 854; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

```
Qy    1 ATGGCCGGCCTCGGCCACCCTCCGCCTTCGGCCGGCCACCCACGCCGTGGTCGGCGGCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368 ATGGCCGGCCTCGGCCACCCTCCGCCTTCGGCCGGCCACCCACGCCGTGGTCGGCGGCG 427

Qy   61 CTACCCGAGTCGCTCGCCAGCACGCGCTGAGAAGCGCCAAGGCGCGAGGAGGTGGACGTC 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  428 CTACCCGAGTCGCTCGCCAGCACGCGCTGAGAAGCGCCAAGGCGCGAGGAGGTGGACGTC 487

Qy  121 GCCCGGCCGCGGAACGGCAGCAGCCAGCTCTACGTGCGTGCTGGGCAGCAAGCTGGGCTG 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488 GCCCGGCCGCGGAACGGCAGCAGCCAGCTCTACGTGCGTGCTGGGCAGCAAGCTGGGCTG 547

Qy  181 CAGGTGGTGGAGCTGCCGGACGAGAGCCTTCCGGACTGCGTTCGTGCGGAGGACGTG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548 CAGGTGGTGGAGCTGCCGGACGAGAGCCTTCCGGACTGCGTTCGTGCGGAGGACGTG 607

Qy  241 GCCGTGGTGTGCGAGGAGGACGGCCCTCATCACCCGACCCCGGGCGGCCGAGCCGAGGAAG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  608 GCCGTGGTGTGCGAGGAGGACGGCCCTCATCACCCGACCCCGGGCGGCCGAGCCGAGGAAG 667

Qy  301 GAGGTTGACATGATGAAAAGAAGCATTAGAAAAACTTCAGTCAATATAGTAGAGATGAAA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668 GAGGTTGACATGATGAAAAGAAGCATTAGAAAAACTTCAGTCAATATAGTAGAGATGAAA 727

Qy  361 GATGAAAATGCAACTTAGATGGGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTG 420
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  728 GATGAAAATGCAACTTAGATGGGGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTG 787

Qy  421 GGCCTTTCCAAAAGACAAATCACCGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788 GGCCTTTCCAAAAGACAAATCACCGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 847
```

```
Qy    481  TATGCAGTCTCCACAGTGCCAGTGCGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848  TATGCAGTCTCCACAGTGCCAGTGCGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATG  907

Qy    541  GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGGCCCTTAAGATC  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGGCCCTTAAGATC  967

Qy    601  ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA  1027

Qy    661  AACTGTATATATCTAAATATCCCCACAAAGGGCACGTCTCTGCTGCCACCGAACCCCGGAA  720
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    1028 AACTGTATATATCTAAATATCCCCACAAAGGGCACGTCTTGCTGCCACCGAACCCCGGAA   1087

Qy    721  GAGTATCCAGAAAGTGCAAAGGTTTATGAAGGACCATATGCTGATCCCGTG           780
           |||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    1088 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTG    1147

Qy    781  AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1148 AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC  1207

Qy    841  AAGAAGTAGACTCCTGA    858
           |||||| ||||||||||
Db    1208 AAGAAAGTAGACTCCTGA  1225
```

RESULT 7

ADE71463

ID   ADE71463 standard; DNA; 858 BP.
XX
AC   ADE71463;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human Met87 variant dimethylarginine dimethylaminohydrolase-1 DNA.
XX
KW   Human; variant dimethylarginine dimethylaminohydrolase-1; variant DDAH 1;
KW   chromosome 1; single nucleotide polymorphism; SNP; ADMA;
KW   asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
KW   coronary heart disease; cerebrovascular disease; hypertension; diabetes;
KW   susceptibility; genotyping; therapeutic targeting; antidiabetic;
KW   cardiant; cerebroprotective; hypotensive; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..858
FT   /*tag= a
FT   /product= "Met87 variant DDAH 1"
FT   replace(260,C)
FT   /*tag= b
FT   variation
FT   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   /note= "Results in Thr (C variant) or Met (T variant) at
FT   position 87 of the protein"
XX
PN   WO2003089638-A1.
XX
PD   30-OCT-2003.
XX
PF   11-APR-2003; 2003WO-FI000274.
XX
PR   19-APR-2002; 2002US-00125456.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Valkonen V, Salonen JT, Pirskanen M, Tuomainen T, Laakso J;
PI   Laaksonen R;
XX
DR   WPI; 2003-854121/79.

DR   P-PSDB; ADE71464.
XX
PT   New nucleic acid encoding a variant dimethylarginine
PT   dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
PT   and its vascular complications, e.g. coronary or cerebrovascular disease
PT   or hypertension.
XX
PS   Claim 4; SEQ ID NO 1; 37pp; English.
XX
CC   The invention relates to a variant dimethylarginine
CC   dimethylaminohydrolase-1 (DDAH 1; ADE71464) and a nucleic acid encoding
CC   it (ADE71463). DDAHs regulate the metabolism of ADMA (asymmetrical
CC   dimethyl-arginine; a naturally occurring nitric oxide synthase inhibitor)
CC   and nitric oxide, which are relevant in disorders such as cardiovascular
CC   disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
CC   has been implicated in susceptibility to diseases such as familial
CC   combined hyperlipidaemia, premature coronary artery disease (CAD), non
CC   insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
CC   The variant DDAH 1 of the invention has Met at position 87, rather than
CC   Thr, which is caused by a C to T change in exon 1 at a position
CC   corresponding to base 260 in ADE71463 and ADE71465. Individuals hetero-
CC   or homozygous for the Thr87Met allele have an increased risk of
CC   developing cardiovascular diseases and diabetes. The invention also
CC   relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a
CC   susceptibility to a cardiovascular disease and diabetes in an individual
CC   by determining the DDAH 1 genotype; a method of treating diabetes or
CC   vascular complications of diabetes by enhancing nitric oxide
CC   availability, production or concentration; a method of targeting
CC   treatment of cardiovascular disease and diabetes in a hypertensive
CC   patient by determining the DDAH 1 genotype and treating them with a drug
CC   that affects nitric oxide availability, production or metabolism; and a
CC   transgenic animal which carries a human variant DDAH 1 nucleic acid
CC   sequence. The nucleic acid molecules and polypeptides are useful for
CC   treating cardiovascular disease (e.g., coronary heart disease,
CC   cerebrovascular disease and hypertension), and diabetes and its vascular
CC   complications. The methods are useful for determining whether a patient
CC   will benefit from treatment with a drug which affects nitric oxide
CC   availability, production or metabolism; a drug which reduces ADMA
CC   availability or concentration; orn an agent which elevates DDAH
CC   availability or concentration (such as DDAH agonist). The methods are
CC   also useful for determining whether a patient will be at risk of adverse
CC   effects if DDAH antagonists are administered. The present sequence
CC   represents a specifically claimed nucleic acid encoding the variant
CC   (Met87) DDAH 1.
XX
SQ   Sequence 858 BP; 218 A; 212 C; 258 G; 170 T; 0 U; 0 Other;

Query Match   99.1%; Score 850; DB 9; Length 858;
Best Local Similarity 99.4%; Pred. No. 9.9e-184;
Matches 853; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      1  ATGGCGGCCTCGGCACACCCTCGGCCCTTCGGCCGGCCACCCACGCCGTGGTGCGGGCG   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGGCGGCCTCGGCACACCCTCGGCCCTTCGGCCGGCCACCCACGCCGTGGTGCGGGCG   60

Qy     61  CTACCCGAGTCGCTCTGCCAGCACGCGCGTGAGAAGCGCCAAGGGCGCGAGGAGGTGACGTC  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CTACCCGAGTCGCTCTGCCAGCACGCGCGTGAGAAGCGCCAAGGGCGCGAGGAGGTGACGTC  120

Qy    121  GCCCGGCCGGGAACGGCAGCAGCCAGTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GCCCGGCCGGGAACGGCAGCAGCCAGTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG  180

Qy    181  CAGGTGGTGGAGCTGCCGACCACGAGAGCCTTCGGACTGGCGTCTTCCGTCGGCCTG      240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CAGGTGGTGGAGCTGCCGACCACGAGAGCCTTCGGACTGGCGTCTTCCGTCGGACGTG     240

Qy    241  GCCGTGGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGCGCCGAGCCGGACGGAGGAAG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GCCGTGGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGCGCCGAGCCGGACGGAGGAAG  300

Qy    301  GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
PR   09-MAY-2002;  2002US-0378653P.
PR   09-MAY-2002;  2002US-0378665P.
PR   09-JUL-2002;  2002US-0394230P.
PR   09-JUL-2002;  2002US-0394253P.
PR   04-SEP-2002;  2002US-0407688P.
PR   28-JAN-2003;  2003US-0442900P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI   Elashoff M;
XX
DR   WPI; 2003-731472/69.
XX
PT   Determining if a compound induces a toxic effect on a tissue or cell, for
PT   identifying hepatotoxic compounds, comprises comparing a gene expression
PT   profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT   mean values.
XX
PS   Claim 44; SEQ ID NO 3556; 874pp; English.
XX
CC   The present invention describes a method for determining whether a
CC   compound induces a toxic effect on a tissue or cell. The method comprises
CC   preparing a gene expression profile of a tissue or cell sample exposed to
CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;

Query Match        83.4%;   Score 715.6;   DB 9;   Length 3008;
Best Local Similarity   89.6%;   Pred. No. 5.5e-153;
Matches   769;   Conservative   0;   Mismatches   89;   Indels   0;   Gaps   0;

Qy     1 ATGGCCGGCTTCGGCCACCCTCCGCCTTCGGCCGGCCACCCCACGCCGTGGTGCGGGCG 60
         |||||||||||| |||||| |||| ||| ||||||||||||||||||||||||||||||  |
Db   432 ATGGCCGGCTTCAGCCACCCCCTCGGTTCTTCGGCCGGCCACCCCACGCCGTGGTGCGGGGCT 491

Qy    61 CTACCCAGTCGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGGCCAAGGGCGGCTGGGCGTGGACGTC 120
         | ||||||||| |||| ||||||||||||||||| ||||  || |    ||| |||||| ||||||| ||||
Db   492 CCGCCCAGTCCCTGCGCCGGCACGCGCTGAGAGCGCCGGGCCTCCCAGGCGCGAAGAGGTGGATTTC 551

Qy   121 GCCCGGCGGAACGGCAGCACCAGTCTCAGTGGCGTGCTCTAGGTGGCAGCAAGCTGGGCTG 180
         ||| | |||||| ||||||||||||| |||| ||||||| |  |||||||||||||||||||
Db   552 GCTCGCCGCTGAGCGCGTGCTGAGCGCGTCTACCGTGGGCGTGCTGGGCAGCAAGCTGGGCTG 611

Qy   181 CAGGTGGTGGAGCGTCGGCGGCGACGAGGAGCCTTCGGACTGCTGTCTTCGGTGGAGGACGTG 240
         ||||||||| ||| | |||||||||||| |||| || |||||||||  ||| |||||||||||
Db   612 CAGGTGGTGCAGCTGCGCGGCGACGAGAGCCTGCTGCACTGCTGTCGTGTTCGTGGAGGACGTG 671

Qy   241 GCCGTTGGTGTGCGAGGAGAGAGGCCCTCATCACCCGACCCGGGCGCCCGGGGGCCTAGCCGCAGGAAG 300
         |||||  |||||||||||||||||||||||||||| | || | ||| ||||| ||| |||||||||||||
Db   672 GCCGTGGTGTGCGAGGAGAGAGGCCCTCATCACCCATCACCCGCCCCGGGCGCGGGCGCCCGGGCTAGCCGCGCAGGAAG 731

Qy   301 GAGGTTGACATGATGAAGAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
         ||||||||||||||||||||| |||| ||| |||||||||||||| |||||||||||||||
Db   732 GAGGTTGACATGATGAAGAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAA 791

Qy   361 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 420
         ||||||||||||||||||||| |||   |  |  |||||||||||||| | ||||||||
Db   792 GATGAAAATGCAACTTAGATGGTGGGGGACCTCCATTCACAGGCAGAGAGAGTTTTTTGTG 851

Qy   421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
         ||||||||||||||||||||||||||||||||||| |||||||||||||||| | |||||
Db   852 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGAC 911
```

```
Db   301 GAGGTTGACATGATGAAGAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
Qy   361 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 420
Qy   421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
Qy   481 TATGCAGTCTCCACAGTGCCAGTGCCAGTGCCGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TATGCAGTCTCCACAGTGCCAGTGCCAGTGCCGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540
Qy   541 GCTGGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAGATC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCTGGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAGATC 600
Qy   601 ATGCAACAGATGAGTGACCACGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATGCAACAGATGAGTGACCACGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
Qy   661 AACTGTATATATCTAAATATCCCCAACAAAGGGCACGTCTTGCTCACCGAACCCCGGAA 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AACTGTATATATCTAAATATCCCCAACAAAGGGCACGTCTTGCTCACCGAACCCCGGAA 720
Qy   721 GAGTATCCAGAAAGTGCAAAGGTTTATGAAGAACTGAAGGACCCATATGCTGATCCCCGTG 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGTATCCAGAAAGTGCAAAGGTTTATGAAGAACTGAAGGACCCATATGCTGATCCCCGTG 780
Qy   781 AGCATGTCTGAACTGGGAAAAGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGCATGTCTGAACTGGGAAAAGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC 840
Qy   841 AAGAAGGTAGACTCCTGA 858
         |||||| |||||||||||
Db   841 AAGAAAGTAGACTCCTGA 858
```

```
RESULT 8
ADB53014
ID   ADB53014 standard; DNA; 3008 BP.
XX
AC   ADB53014;
XX
DT   04-DEC-2003 (first entry)
XX
DE   Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3556.
XX
KW   toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW   toxicity marker; toxicity progression; drug screening;
KW   primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO2003065993-A2.
XX
PD   14-AUG-2003.
XX
PF   04-FEB-2003; 2003WO-US003482.
XX
PR   04-FEB-2002;  2002US-0353171P.
PR   13-MAR-2002;  2002US-0363534P.
PR   08-APR-2002;  2002US-0370248P.
PR   10-APR-2002;  2002US-0371134P.
PR   10-APR-2002;  2002US-0371135P.
PR   10-APR-2002;  2002US-0371150P.
PR   11-APR-2002;  2002US-0371413P.
PR   19-APR-2002;  2002US-0373601P.
PR   19-APR-2002;  2002US-0373602P.
PR   22-APR-2002;  2002US-0374139P.
PR   08-MAY-2002;  2002US-0378370P.
PR   09-MAY-2002;  2002US-0378652P.
```

Right column:

```
SQ     Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;

Query Match          83.4%;  Score 715.6; DB 9;  Length 3008;
Best Local Similarity 89.6%;  Pred. No. 5.5e-153;
Matches 769;  Conservative 0;  Mismatches 89;  Indels 0;  Gaps 0;

Qy     1 ATGGCCGGCTCGGCCACCCTCCGGCCTTCGGCCGGCGGCCACCCGCCGTGGTGCGGGCG   60
         ||||||||||  |  |  ||||| ||||||||||||||||||||||||||||||||||||
Db   432 ATGGCCGGCCTCAGCTCACCCCTCGGTCTTCGGCCGGCGGCCACCCGCCGTGGTGCGGGCt 491

Qy    61 CTACCCAGTCGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC  120
         |||    |||||  ||||||||| || |||||  ||||||||||||||||||||||||||
Db   492 CCGGCCCGAGTCCCTGCGCCAGCGCGTCGAGCGCGTTCCAGGGCGAAGAGGTGGATTTC 551

Qy   121 GCCCGGCGGAACGGCAGCAGCCAGCTTACGTGGGCGTGCTGGGCGCAAGCAAGCTGGGCTG  180
         ||||| ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   552 GCTCGGCTGAGCGGCCAGCACCAGCTTACGTGGGCGTGGGCAGCAAGCTGGGCGTG 611

Qy   181 CAGTGGTGGAGCTGCCGGCGACGAGCCCTTCCGGACTGCGTCTTCGTGGAGGACGTG  240
         ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   612 CAGTGGTGGCAGCTGCCGGCGACGAGCCTGCCGCCTGCGTGTTCGTGGAGGACGTG 671

Qy   241 GCCGTGGTGTGCGAGGAGACGACGGCCCTCATCACCCGCGGCGGCGCCAGCCGGACAGGAAG  300
         ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   672 GCCGTGGTGTGCGAGGAGACGACCCCTCATCACCCGCGGGCGGCGCCTAGCCCGCAGGAAG 731

Qy   301 GAGGTTGACATGATGAAAGAAGCATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   732 GAGGTTGACATGATGAAAGAAGAGCTTTGAAAAACTTCAGCTCAACATAGTAGAGATGAAA 791

Qy   361 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   792 GATGAAAATGCAACTTTAGATGTGGGGGACGTCCTATTCACAGGCAGAGAGTTTTTTGTG 851

Qy   421 GGCCTTCCAAAAGGACAAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   852 GGCCTTTCCAAAAGGACAAAATCAACGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGAC 911

Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG  540
         |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   912 TACGCAGTTTCCACAGTCCCCGTGGGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCATG 971

Qy   541 GCTGGGCCTAACCTGATCGCAATTGGTCTAGTGAATCTGCACAGAAGGCCCTAAGATC  600
         |||||| ||||||||||||||||||||||||||||||||| |||||||||||| |||||
Db   972 GCTGCCCCAACCTGATCGCATAGGGTCCAGTGAATCTGCGCAGAAGGCCCTCAAGATC 1031

Qy   601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGCCTGATGACATAGCAGCA  660
         |||||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db  1032 ATGCAACAGATGAGTGACCACCGTTATGCAACAAGCTCACTGGACGACATGGCCGCC 1091

Qy   661 AACTGTATATATCTAAAATATCCCAACAAAGGCACGTCTTGCTGCCACCGAACCCCGAA  720
         |||||||||||| ||||||||||| ||||||||| |||||||||||||||||||||||| 
Db  1092 AACTGTATATATTTAAATATCCCAGCAAAGGCATGTCTTGCTGCCACCGAACCCCAGAA 1151

Qy   721 GAGTATCCAGAAAGTGCAAAGGTTATGAGAAACTGAAGGACCATATGCTGATCCCGTG  780
         ||||| |||||||| ||||||||| |||||| || ||||||||| | |||||||||||
Db  1152 GAGTACCCAGAAAGCGCAAAGGTTTATGAGAAGCTCAAGGACCATCTACTGATCCCGTG 1211

Qy   781 AGCATGTCTGAACTGGAAAAGGTGATGGGCTGCTCACCTGCTCCTCAGTTTAATTAAC  840
         |||| | |||| ||||||||||||||| ||||  |||||||||| ||  |||||||||
Db  1212 AGCAATTCTGAGATGGAAAAGGTGACGGGCTTGCTCACCTGCTCCGTTTTTATTAAC 1271

Qy   841 AAGAAGGTAGACTCCTGA  858
         |||||  ||||| ||||||
Db  1272 AAGAGACAGACTCTTGA 1289

RESULT 10
ADE53981
ID   ADE53981 standard; cDNA; 3778 BP.
XX
AC   ADE53981;
XX
```

Left column (continuation):

```
Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG  540
         |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   912 TACGCAGTTTCCACAGTCCCCGTGGGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCATG 971

Qy   541 GCTGGGCCTAACCTGATCGCAATTGGTCTAGTGAATCTGCACAGAAGGCCCTAAGATC  600
         |||||| ||||||||||||||||||||||||||||||||| |||||||||||| |||||
Db   972 GCTGCCCCAACCTGATCGCATAGGGTCCAGTGAATCTGCGCAGAAGGCCCTCAAGATC 1031

Qy   601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGCCTGATGACATAGCAGCA  660
         |||||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db  1032 ATGCAACAGATGAGTGACCACCGTTATGCAACAAGCTCACTGGACGACATGGCCGCC 1091

Qy   661 AACTGTATATATCTAAAATATCCCAACAAAGGCACGTCTTGCTGCCACCGAACCCCGAA  720
         |||||||||||| ||||||||||| ||||||||| ||||||||||||||||||||||||
Db  1092 AACTGTATATATTTAAATATCCCAGCAAAGGCATGTCTTGCTGCCACCGAACCCCAGAA 1151

Qy   721 GAGTATCCAGAAAGTGCAAAGGTTATGAGAAACTGAAGGACCATATGCTGATCCCGTG  780
         ||||| |||||||| ||||||||| |||||| || ||||||||| | |||||||||||
Db  1152 GAGTACCCAGAAAGCGCAAAGGTTTATGAGAAGCTCAAGGACCATCTACTGATCCCGTG 1211

Qy   781 AGCATGTCTGAACTGGAAAAGGTGATGGGCTGCTCACCTGCTCCTCAGTTTAATTAAC  840
         |||| | |||| ||||||||||||||| ||||  |||||||||| ||  |||||||||
Db  1212 AGCAATTCTGAGATGGAAAAGGTGACGGGCTTGCTCACCTGCTCCGTTTTTATTAAC 1271

Qy   841 AAGAAGGTAGACTCCTGA  858
         |||||  ||||| ||||||
Db  1272 AAGAGACAGACTCTTGA 1289

RESULT 9
ADB79843
ID   ADB79843 standard; DNA; 3008 BP.
XX
AC   ADB79843;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Rat N-G,N-G-dimethylarginine dimethylaminohydrolase DNA, SEQ ID 83.
XX
KW   Analgesic; pain; streptozocin-induced diabetes; rat; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   EP1279744-A2.
XX
PD   29-JAN-2003.
XX
PF   26-JUL-2002; 2002EP-00255249.
XX
PR   27-JUL-2001; 2001GB-00018354.
PR   07-FEB-2002; 2002GB-00002910.
XX
PA   (WARN ) WARNER LAMBERT CO.
XX
PI   Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
XX
DR   WPI; 2003-395407/38.
DR   P-PSDB; ADB79842.
XX
PT   Use of isolated gene sequences and encoded polypeptides that are
PT   upregulated in the spinal cord in response to streptozocin-induced
PT   diabetes for screening compounds for the treatment of pain, or for
PT   diagnosing pain.
XX
PS   Claim 1; Page 155; 334pp; English.
XX
CC   The present invention relates to nucleotide sequences which are useful in
CC   the screening of compounds for the treatment of pain, or for the
CC   diagnosis of pain. The nucleotide sequences are up-regulated in the
CC   spinal cord in response to streptozocin-induced diabetes. The present
CC   sequence is one such nucleotide sequence.
XX
```

```
Db   301  TCCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGC  360

Qy   500  CAGTGGCGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCG  559
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CAGTGGCGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCG  420

Qy   560  CAATTGGGCTCAGTGAAATCTGCACAGAGAGGCCCTAAGATCATCATCGCAACAGATGAGTGACC  619
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CAATTGGGCTCAGTGAAATCTGCACAGAGAGGCCCTAAGATCATCGCAACAGATGAGTGACC  480

Qy   620  ACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATA  679
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  ACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATA  540

Qy   680  TCCCCAACAAAGGGCACGTCTTGTGCACCGACCCCGGAAGAGTATCCAGAAAGTGCAA  739
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  TCCCCAACAAAGGGCACGTCTTGTGCACCGACCCCGGAAGAGTATCCAGAAAGTGCAA  600

Qy   740  AGGTTTATGAGAAACTGAAGGACCATATGTCTGATCCCCGTGAGCATGTCTGAACTGGAAA  799
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  AGGTTTATGAGAAACTGAAGGACCATATGTCTGATCCCCGTGAGCATGTCTGAACTGGAAA  660

Qy   800  AGGTGGATGGGCTGCTCACCTGCTCCAGTTTTAATTAACAAGAAGTAGACTTCCTGA  858
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  AGGTGGATGGGCTGCTCACCTGCTCCAGTTTTTAATTAACAAGAAGTAGACTTCCTGA  719
```

RESULT 11
AAFI6077
ID   AAFI6077 standard; cDNA; 3651 BP.
XX
AC   AAFI6077;
XX
DT   13-MAR-2001 (first entry)
XX
DE   Human prostate cancer antigen nucleotide sequence SEQ ID NO:512.
XX
KW   Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW   neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW   vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW   antibacterial; gene therapy; neural; immune; reproductive; renal;
KW   gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW   wound; infectious disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055174-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005988.
XX
PR   12-MAR-1999; 99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2000-587513/55.
DR   P-PSDB; AAB56874.
XX
PT   Prostate cancer associated gene sequences, referred to as prostate cancer
PT   antigens, useful for treatment, prevention, and diagnosis of disorders
PT   such as prostate cancer.
XX
PS   Claim 1; Page 976-977; 2338pp; English.
XX
CC   AAFI5566 to AAFI6505 encode the human prostate cancer associated
CC   proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC   The prostate cancer antigens can have neuroprotective, cytostatic,
CC   cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC   nephrotropic, antiinfective, gynaecological and antibacterial activities,
```

---

```
DT   29-JAN-2004  (first entry)
XX
DE   Human prostate cancer cDNA #328.
XX
KW   Human; prostate cancer; ss; cDNA combination; differential expression;
KW   gene.
XX
OS   Homo sapiens.
XX
PN   US2003190640-A1.
XX
PD   09-OCT-2003.
XX
PF   29-MAY-2002; 2002US-00252157.
XX
PR   31-MAY-2001; 2001US-0295048P.
XX
PA   (FARI/) FARIS M.
PA   (PEAR/) PEARSON C I.
XX
PI   Faris M,  Pearson CI;
XX
DR   WPI; 2003-831619/77.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   prostate cancer, useful for diagnosing, treating or monitoring the
PT   progression of treatment of prostate cancer.
XX
PS   Claim 1; SEQ ID NO 328; 42pp; English.
XX
CC   The invention relates to a combination comprising a number of cDNAs
CC   expressed in prostate cancer. The invention also relates to a method for
CC   detecting differential expression of one or more cDNAs in a sample
CC   containing nucleic acids by hybridising a substrate with the nucleic
CC   acids, thus forming one or more hybridisation complexes, detecting
CC   hybridisation complex formation and comparing the complexes formed with
CC   standard complexes, where differences between the standard and the sample
CC   complex formation indicate differential expression of cDNAs in the
CC   sample. The differential expression is diagnostic of prostate cancer. The
CC   invention also relates to proteins and antibodies related to the cDNAs.
CC   The combination is useful for diagnosing, treating or monitoring the
CC   progression of treatment of prostate cancer. The antibodies are useful
CC   for detecting prostate cancer. This sequence represents a human prostate
CC   cancer cDNA of the invention.
XX
SQ   Sequence 3778 BP; 1064 A; 821 C; 719 G; 1150 T; 0 U; 24 Other;

Query Match     82.2%;  Score 705.4;  DB 9;  Length 3778;
Best Local Similarity 99.7%;  Pred. No. 1.2e-150;
Matches  717;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy   141  CCAGCTCTACGTGGGCCGTGCTGGGCAGCAACTGGGGCTCCAGGTGGTGAGCTGCCGGC  200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  CCAGCTCTACGTGGGCCGTGCTGGGCAGCAACTGGGGCTCCAGGTGGTGAGCTGCCGGC  60

Qy   201  CGACGAGGAGAGCCTTCCGGACTGCCTCTTCGTGGAGACACGTGGCCGTGGTGTGCGGAGGAGAC  260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CGACGAGGAGAGCCTTCCGGACTGCCTCTTCGTGGAGACACGTGGCCGTGGTGTGCGGAGGAGAC  120

Qy   261  GGCCCTCATCACCCGACACCCGGGGCGGGGGCCGAGCCCGGAGGAGGTTGACATGATGAAAGA  320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GGCCCTCATCACCCGACACCCGGGGCGGGGGCCGAGCCCGGAGGAGGTTGACATGATGAAAGA  180

Qy   321  AGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAGAAGATGAAAGATGAAACTTTAGA  380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  AGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAGAAGATGAAAGATGAAACTTTAGA  240

Qy   381  TGGCCGGAGATGTTTATTCACAGACGCAGAGAATTTTTGTGGGCCTTTCAAAAGGACAAA  440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TGGCCGGAGATGTTTTATTCACAGACGCAGAGAATTTTTGTGGGCCTTTCAAAAGGACAAA  300

Qy   441  T-CAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGC  499
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

CC   and can be used in gene therapy. The prostate cancer antigen
CC   polynucleotides may be used for detection of prostate cancer, chromosome
CC   identification, as chromosome markers, and for numerous other diagnostic
CC   or research purposes. The prostate cancer antigens may be used to treat
CC   disorders such as neural, immune, muscular, reproductive,
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC   AAB57303 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 3651 BP; 1016 A; 801 C; 684 G; 1141 T; 0 U; 9 Other;

Query Match   73.5%; Score 630.4; DB 3; Length 3651;
Best Local Similarity   99.7%; Pred. No. 1.4e-133;
Matches 642; Conservative   0; Mismatches   0; Indels   1; Gaps   1;

Qy   215 CGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACGGCCCTCATCACCC 274
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 CGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACGGCCCTCATCACCC 61

Qy   275 GACCCGGGGCGGCCCAGCCCGGAGGAGGAGGAGTTGACATGATGAAAGAAGCATTAGAAAAAC 334
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 GACCCGGGGCGGCCCAGCCCGGAGGAGGAGGAGTTGACATGATGAAAGAAGCATTAGAAAAAC 121

Qy   335 TTCAGCTCAATATAGTAGAGATGAAAGATGAAAAATGCAACTTTAGATGGCGGAGATGTTT 394
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 TTCAGCTCAATATAGTAGAGATGAAAGATGAAAAATGCAACTTTAGATGGCGGAGATGTTT 181

Qy   395 TATTCACAGGCAGAGAATTTTTGTGGGCCTTCCAAAAGGACAAATCAACGAGGTGCTG 454
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 TATTCACAGGCAGAGAATTTTTGTGGGCCTTCCAAAAGGACAAATCAACGAGGTGCTG 241

Qy   455 AAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGGT 514
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 AAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGGT 301

Qy   515 TGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTCAGTG 574
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 TGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTCAGTG 361

Qy   575 AATCTGCACAGAAGGCCCTTAAGATCATGCAACAGAGAGTGACCACCGCTACGACAAAC 634
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 AATCTGCACAGAAGGCCCTTAAGATCATGCAACAGAGAGTGACCACCGCTACGACAAAC 421

Qy   635 TCACTGTGCCTGATGACATAGCAGCAGAGTATCCCAACAGTATATATCTAAATATCCCAACAAAGGGC 694
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 TCACTGTGCCTGATGACATAGCAGCAGAGTATCCCAACAGTATATATCTAAATATCCCAACAAAGGGC 481

Qy   695 ACGTCTTGCTGCACCGGAACCCCGGAAGAGTATCCAGAAGGTTTATGAGAAAC 754
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 ACGTCTTGCTGCACCGGAACCCCGGAAGAGTATCCAGAAGGTTTATGAGAAAC 541

Qy   755 TGAAGGACCATATGCTGATCCCCGTGAGCATGTCGAACTGGATGGGCTGC 814
               |||||||||||||||||||||||||||||||||||||||||||||
Db   542 TGAAGGACCATATGCTGATCCCCGTGAGCATGTCGAACTGGATGGGCTGC 601

Qy   815 TCACCTGCTGCTCAGTTTTAATTAACCAAGAAGTAGACTCCTGA 858
               ||||||||||||||| ||||||||||||||||||||||||||||
Db   602 TCACCTGCTGCTG-TCAGTTTTAATTAACCAAGAAGTAGACTCCTGA 644

RESULT 12
ABN89945
ID   ABN89945 standard; cDNA; 444 BP.
XX
AC   ABN89945;
XX
DT   16-AUG-2002 (first entry)
XX
DE   Mouse clone DST IMX3_103 cDNA.
XX
KW   Mouse; antiinflammatory; gene therapy; ileitis; DST; ss; TOGA;
KW   digital sequence tag; total gene expression analysis.

XX
OS   Mus musculus.
XX
PN   WO200231114-A2.
XX
PD   18-APR-2002.
XX
PF   11-OCT-2001; 2001WO-US032091.
XX
PR   11-OCT-2000; 2000US-0239483P.
XX
PA   (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
XX
PI   Viney JL, Sims JE, Dubose RF, Baum PR, Hasel KW, Hilbush BS;
XX
DR   WPI; 2002-426279/45.
XX
PT   New isolated nucleic acid molecules that are associated with ileitis, for
PT   preventing, treating, modulating and diagnosing ileitis in a mammalian
PT   subject.
XX
PS   Claim 1; Page 212-213; 273pp; English.
XX
CC   The invention relates to a novel isolated nucleic acid molecule
CC   comprising a polynucleotide having one of 90 polynucleotide sequences,
CC   given in the specification. The polynucleotides of the invention have
CC   antiinflammatory activity, and may have a use in gene therapy. The
CC   polynucleotide or a polypeptide encoded by it is used for preventing,
CC   treating, modulating or ameliorating a medical condition such as ileitis.
CC   The polypeptide or polynucleotide is also useful for manufacturing a
CC   medicament for treating ileitis. The sequence represents a cDNA digital
CC   sequence tag obtained from a mouse clone by the TOGA (total gene
CC   expression analysis) method
XX
SQ   Sequence 444 BP; 119 A; 126 C; 101 G; 98 T; 0 U; 0 Other;

Query Match   25.5%; Score 218.8; DB 6; Length 444;
Best Local Similarity   82.7%; Pred. No. 4e-40;
Matches 263; Conservative   0; Mismatches   52; Indels   3; Gaps   1;

Qy   544 GGGCCTAACCTGATCGCAATTGGGTCTCAGTGAATCTGCACAGAAGGCCCTTAAGATCATG 603
                ||||||||||||| |||||| ||||| ||| ||||||||||||||| |||||||||||||
Db   2 GGACCCAACCTGATTGCAATAGGGTCCCAGCGAATCTGCACAGAAGAAGGCCCTCAAGATCATG 61

Qy   604 CAACAGATGAGTGACCACCGC--TACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
               ||||||||||| ||| ||||| ||||||||| |||||||||||||||||||||||||
Db   62 CAACAGATGAAGTGCCCCATCGTTAATGCCAGTTCCCGTACCGGACGACATGGCCGCCC 121

Qy   661 AACTGTATATCTAAATATCCCAACAGTATATCCCCAACAAAGGGCACGTCTTGCTGCACCGGAACCCCGAA 720
               ||||| ||||||||||||||| ||||| |||||||||||||||||||||||||||||||
Db   122 AACTGCATATCTAAATATCTCAACAGTCTATGCCAACAAAGGGCATGTCTTGCTGCACCGACCCCAGAA 181

Qy   721 GAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAACTGCTGGATCCCCTG 780
               ||||| ||||||||||||| |||||||||||||| |||||| |||||||||||
Db   182 GAGTACCCAGAAAGCGCAAAGTCTATGAGAAACTCAAGGAGCCATCCGATCCCTGTG 241

Qy   781 AGCATGTCGAACTGGAAAAGGTGGATGGGCTGTGCTCACCTGCTCACGTGCTCAGTTTTAATTAAC 840
               |||||| || ||||| ||||| |||| ||| ||||||||||||||||||||||| ||||||||||
Db   242 AGCAACTCGGAGATGGGAAAAGGCGGAAAAGGTGGACGCGTCGCCTCCTCACGTGCTCCGTTTTTTATTAAC 301

Qy   841 AAGAAGGTAGACTCCTGA 858
               |||||||| |||||||||
Db   302 AAGAAGATAGACTCCTGA 319

RESULT 13
AAZ52967
ID   AAZ52967 standard; cDNA; 1376 BP.
XX
AC   AAZ52967;
XX
DT   14-MAR-2000 (first entry)
XX

```
Qy   509  ATGGGTTGCATTGAAGAGTTTCTGCAGCATGCTGGGCCTCAACCTGATCGCAATTGGGT  568
          ||||||||||||||||||||||||||||||||| |||||||  |||||||| ||||||||
Db   799  GTCCCTCCCACCTGCGCGGGTCTGACGTCTGGGCATGGGGGACCTCGCACTGTTGTGGCAGGCA  858

Qy   569  CTAGTGAATCTGCACAGAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTACG  628
          |||   ||||| |||||||| ||||  |||||  |||||||||  ||||||||||||||||||
Db   859  GCAGCGACGCGCTGCCCAAAAGGCTGTCCGGGCAATGGCAGTGCTGACAGATCACCCATATG  918

Qy   629  ACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGATATATCTAAT-----ATCC  682
          | ||||||| ||||||||||| || ||| || || ||| ||||||||||
Db   919  CCTCCCTGACCCTCCAGATGACGCAGCTGCTGACGTCGTCTTTCTTCGTCCTCTGGGTTGC  978

Qy   683  CCAACAAAGGGCACGTCTTGCTGCACCGAACCCGGAAGAGTATCCAGAAGTGCAAAGG  742
          |||||||||||||   ||||||||||||||||||||||||||||
Db   979  CTGGTGTGCCCCCTTTCCTCCTGCACCGTGAGGTGGGATCTGCCCAACAGCCAGGAGG  1038

Qy   743  TTTATGAGAAACTGAAGGACCATATGCTGATCCCGAGCATGTCTGAACTGGAAAAGG  802
          |||  ||||| |||||||  ||||| |||||| |||||||||||||||||||||||||
Db   1039 CACTGCAGAAGTCTCTTGATGTGACCCTGGTACCCTGTTCCTGGTGTCCTCAGAACTGGAAGAGG  1098

Qy   803  TGGATGGGCTGCTCACCTGCTCAGTTTAATTAACA  841
          || |||||| ||||||| |||  ||||
Db   1099 CTGGCGCGGGCTCAGCTGCCCCTGCCTGCTGGTCGTCAGCCA  1137
```

```
RESULT 14
AAA47656
ID   AAA47656 standard; cDNA; 858 BP.
XX
AC   AAA47656;
XX
DT   08-NOV-2000  (first entry)
XX
DE   Dimethylarginine dimethylaminohydrolase (DDAH2) coding sequence.
XX
KW   Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH2;
KW   arginine deaminase; hyperlipidemia; renal failure; hypertension;
KW   restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;
KW   ischemia reperfusion injury; septic shock; multi organ failure;
KW   arthritis; skin disorders; inflammatory cardiac disease; migraine;
KW   infection; ds.
XX
OS   Homo sapiens.
XX
FH                     Location/Qualifiers
FT   CDS               1..858
FT                     /*tag= a
FT                     /product= "DDAH2"
XX
PN   WO200044888-A2.
XX
PD   03-AUG-2000.
XX
PF   26-JAN-2000; 2000WO-GB000226.
XX
PR   26-JAN-1999; 99GB-0001705.
PR   04-JUN-1999; 99GB-0013066.
XX
PA   (UNLO ) UNIV COLLEGE LONDON.
XX
PI   Vallance PJT, Leiper JM, Whitley GSJ, Charles IG;
XX
DR   WPI; 2000-543392/49.
DR   P-PSDB; AAB01478.
XX
PT   Novel methylarginase polypeptides and polynucleotides, used to identify
PT   modulators of them, which are used in the treatment of e.g. cancer,
PT   hypertension, and bacterial infections.
XX
PS   Claim 1; Page 57-58; 68pp; English.
XX
CC   Nucleotides encoding methylarginase polypeptides, vectors comprising
CC   these nucleotides and the polypeptides themselves can be used in
```

```
DE   Human prostate tumor cDNA library derived EST fragment #110.
XX
KW   Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
KW   treatment; ds.
XX
OS   Homo sapiens.
XX
PN   DB1982O190-A1.
XX
PD   04-NOV-1999.
XX
PF   28-APR-1998;  98DE-01020190.
XX
PR   28-APR-1998;  98DE-01020190.
XX
PA   (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI   Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR   WPI; 1999-621386/54.
DR   P-PSDB; AAY74144, AAY74145, AAY74146.
XX
PT   New human nucleic acid sequences from pancreatic tumors, and related
PT   proteins.
XX
PS   Claim 2; Page 272; 502pp; German.
XX
CC   This invention describes novel polypeptides and their encoding nucleic
CC   acids derived from human pancreatic tumor tissue which have cytostatic
CC   activity. The sequences are also useful in producing pharmaceutical
CC   compositions for treatment of pancreatic tumors. AAZ52858-Z53014
CC   represent expressed sequence tag (EST) fragments derived from a human
CC   pancreatic tumor cDNA library and which encode the proteins represented
CC   in AAY73814-Y74252
XX
SQ   Sequence 1376 BP; 294 A; 390 C; 448 G; 244 T; 0 U; 0 Other;

Query Match      25.2%;  Score 216.2; DB 2;  Length 1376;
Best Local Similarity  54.9%;  Pred. No. 2.1e-39;
Matches  450;  Conservative  0;  Mismatches  363;  Indels  6;  Gaps  1;
```

```
Qy    29  TCGGCCGGCGCCACCCAGCGCGTGGTGCGGGGCGTACCGAGTCCTGCCAGCACGCGC  88
          ||| |||||||||||||||||||||||||||||||||||||||
Db   319  TGGGCCGCTGCTCCCATGCCGTGGGGGAGTCCCAGAGAGCTGGCGTCGGGGGAG  378

Qy    89  TGAGAAGCAGCCAAGGGCGAGGAGGTGGACGTGGGCGCTGCCGGGCCGGGAACGGCAGCACCAGCTCT  148
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   379  GTGCGGGGGCTGGCCTTCCCGCTCTGGATCTGGCCAAGCTCAAAGGGAGCACGGGGTGC  438

Qy   149  ACGTGGGCGTGCTGGCGCAGCAAGCTGGGGCTGCAGGTGGTGAGCTGTGAGCGCCGACGAGA  208
          ||| ||||||||||||||||||||||||||| ||| |||||
Db   439  TGGGAGGTAAACTGAGGCAACGACTGGGGCTACAGCTGGACTGCTAGAACTGCACCTGAGGAGT  498

Qy   209  GCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACGGCCTCA  268
          |||||||||||||||||||||||||||||||||||||||
Db   499  CATTGCCGCTGGGACCGCTGCTTGGCCGACACGCCGCCGTGATCCAAGCGGCGCCCTAA  558

Qy   269  TCACCCGACCCGGGGCGCGCCGAGCCGAGGAGGAGGTTGACATGATGAAAGAGAGCATTAG  328
          |||||||||| ||||| |||  ||||||||||  ||||||||||||||
Db   559  TCACGCGGCCCTGAGCCCCGGCTCGTAGGCCAGGAGGTCGATGGAGTCCGCAAAGCCCTGC  618

Qy   329  AAAAAACTTCAGCTCAATATAGTAGAGACATGAAGAATGAAAATGAACTTTAGATGGCGGAG  388
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   619  AAGACCTGGGGCTCGAATTGTGAAATAGGAGACGAGAAGACGAGCACGTGGATGGCACTG  678

Qy   389  ATGTTTATTCACGGCAGGAGAATTTTTGTGGGCCTTTCCAAAAGACAAATCAACGAG  448
          | |||  ||||||||||||||||| |||||| |||||||||||||||
Db   679  ACGTTCTCTTCACCGGCCGGGAGTTTTTCGGTAGGCCTCTCCAAATGACCAATCACCGAG  738

Qy   449  GTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCCACAGTGCCAGTGGCAG  508
          ||| |||||||||||||||| | |||||||
Db   739  GAGCTGAGATCGTGCGACACGTTCCGGACTTCGCCGCTCCCGCTCTCCCACTGTGCCGTCTCGG  798
```

CC   medicaments for the treatment of hyperlipidemia, renal failure,
CC   hypertension, restenosis after angioplasty, atherosclerosis,
CC   complications of heart failure, schizophrenia, multiple sclerosis or
CC   cancer. Modulators of the enzyme can be used in medicaments for the
CC   treatment of ischemia-reperfusion injury of the brain or heart, cancer,
CC   lethal hypertension in severe inflammatory conditions such as septic
CC   shock or multi-organ failure, or local and systemic inflammatory
CC   disorders including arthritis, skin disorders, inflammatory cardiac
CC   disease, migraine, or microbial or bacterial infection
XX
SQ   Sequence 858 BP; 150 A; 268 C; 281 G; 159 T; 0 U; 0 Other;

Query Match   25.0%;   Score 214.6; DB 3; Length 858;
Best Local Similarity   54.8%;   Pred. No. 4.3e-39;
Matches 449;   Conservative   0;   Mismatches 364;   Indels   6;   Gaps   1;

Qy   29 TCGGCGGGCGGCCACCCACGCCGTGGTGCGGGCGGCCTACCCGAGTGCTCTGCCAGCACGCGC 88
          ||||||||||||||||||| || | |||||| |||||||| |||||||||| ||
Db   23 TGGGCGGCGCTGCTCCCATGCCTGATCCGGGAGTCCCAGAGAGCCTGGCGTCGGGGGAAG 82

Qy   89 TGAGAAGCCGCCAAGGGCGAGGAGGTGGACGTCGCCGCCGGGAACGGCACGCAGCCAGCTCT 148
          ||||| |||||||||||| ||||| ||||||||||| |||||||||| ||
Db   83 GTGCGGGGGGCTGGCCTTCCCGCTCTGGAAATGCTCGGCCAAAGCTCAAAGGAGCACGGGGTGC 142

Qy   149 ACGTGGCGGTGTCGGGCCAGCAAGCTGGGGCTGCAGGTGCTGTGGAGCTGCCGGCCGACGAGA 208
          |||| ||||||||| ||||||||| ||||| ||||||||||||||||
Db   143 TGGGAGGTAAACTGAGGCAACGACTGGGGCTACAGCTGCTCTAGAACTGCCACCTGAGGAGT 202

Qy   209 GCCTTCCGGACTGCGCTCTTCGTGGAGGACGTGGCCGTGTCGTGGAGGAGACCGCCCTCA 268
          |||||||||||| |||||| ||||||||||||| |||||||| |||||||||||
Db   203 CATTGCCGCTGGACCGCTGCTTGCGCACACGGCCCGTGATCCAAGGGGACACGGCCCTAA 262

Qy   269 TCACCCGACCCGCGGGGGCGGCGCCGAGCCGGAGGAAGAGGTTGACATGATGAAAGAGCATTAG 328
          ||||||||||||||||| | || |||||| |||| |||| ||||||||||||||||
Db   263 TCACCCGGCCCCTGGAGCCCCGCTCGTAGGCCAGAGGTCGATGGAGTCCGCAAAGCCCTGC 322

Qy   329 AAAAACTTCAGCTCAATATAGTAGACAAGATGAAAAATGCAACTTTAGATGGCGGAG 388
          ||| || |||||| ||||| ||||||||||| ||| ||||| | |||||||||||
Db   323 AAGACTGGGGCTCCGAATTGTGGAAATAGGACGAGACGCCGACGCGGACTGGACGACTG 382

Qy   389 ATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAG 448
          |||||||| ||||||||||| ||| ||||||||||| ||||||||||||||||||
Db   383 ACGTTCTCTTCACCCGCGGAGTTTTTCGTAGGCCTCCCAAATGACCAATCACCGAG 442

Qy   449 GTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTCTGCCAGTGGCAG 508
          |||||||||||||||||||||||| ||||||||| |||||||||||| |||| ||||||||
Db   443 GAGCTGAGATCGTGGCGGACACGTTCCGGGACTTCGCCGTCTCCACTGTGTCCCAGTCTCGG 502

Qy   509 ATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCTAACCTGCCAATTGGGT 568
          |||| |||| ||||||||||||||| |||||||||||||||||||||||||
Db   503 GTCCCTCCCACCCGGCCGGTCTCTGCGGCATGGGGCATGGGGGACCTCGCACTGTTGTGGCAGCCA 562

Qy   569 CTAGTGAAATCTGCACAGAGGGCCCTTAAGGATCATGCAACAGATGAGTGACCACCGCTACG 628
          |||||||||||||| ||||||||||| |||| ||||||||||| ||||||| ||||||||
Db   563 GCAGGCACGCTGCCCAAAAGGCTGTCCGGGCAATGGCAGTGCTGCTGCACAGATCACCCATATG 622

Qy   629 ACAAACTCACTGTGCCTGATGACAGTAGCAGCAAACTGATATATCTAAAT------ATCC 682
          |||||| ||||||||| |||| |||||||| || || |||| |||||| ||||
Db   623 CCTCCCTGACCCCTCCAGATGACCAGCAGTCGCTGCACGTGTCTTTCTTCGTCCTGGGTTGC 682

Qy   683 CCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAAGG 742
          |||| || ||||||||||| ||||||||||||| |||||||||| |||||||
Db   683 CTGGTGTGCCCCTTTCCCTCCTGACGTGGAGGTGGGGGATCGGGGGACAGCGCCAGGAGG 742

Qy   743 TTTATGAGAAACTGAAGGACCATATGTCCATTGCTCATCCCCGTGAGCATGTCTGAACTGAAAAGG 802
          ||||||||||||||| |||| ||| || ||| ||||||| |||||||||||||| |||||
Db   743 CACTGCAGAAGCTCTCTGATGTCACCCTGGTACCCTGTGTCCTGCTCAGAAGACCTGAGAAAG 802

Qy   803 TGGATGGCGTGCTCACCTGCTGCCAGTTTAATTAACA 841
          |||||||||||||| ||| |||| ||||||| ||||
Db   803 CTGGCGCCGGGCTCAGCTCCCCTCCCTGCTTGGTGCTCAGCA 841

```
Query Match        25.0%;  Score 214.6;  DB 3;  Length 1228;
Best Local Similarity   54.8%;  Pred. No. 4.7e-39;
Matches 449;  Conservative   0;  Mismatches 364;  Indels   6;  Gaps   1;

Qy      29 TCGGCCGGCCACCCACGCGTGGTGCGGGCGCTACCCGAGTCGCTCTGCCAGCACGCGC      88
           |||||||||||||| ||| |  |||| |  |  |||||| |||  |  | | ||  |||
Db     285 TGGGCCGCTGCTCCCATGCCCTGATCCCGGGGAGTCCCAGAGAGCCTGGCGTCGGGGGAAG  344

Qy      89 TGAGAAGCGCCAAGGGCGAGGAGGTGACGTCGCCGCGCGGAACGGCAGCACCAGCTCT     148
           ||||| | |  |   | ||  | ||| ||| | | | |||  |  | | |||| |
Db     345 GTGCGGGGGGCTGGCCTTCCCGCTCTGGATCTGCCAAAGCTCAAAGGGAGCACGGGGTGC  404

Qy     149 ACGTGGGCGTGCTGGGCAGCAAGCTGGGCTGCAGGTGGTGAGCTGCCGGCCGACGAGA     208
           |  || || | |  |||| ||||| |||  | | || ||| || | | ||| || || |
Db     405 TGGGAGGTAAACTGAGGCAACGACTGGGGCTACAGCTGCTAGAACTGCCACCTGAGGAGT  464

Qy     209 GCCTTCCGACTGGCTCTTCGTGGACGTGGCGTGGTGTGCGAGGAGACGGCCCTCA     268
           |  ||   |  | |  | |  |  | | | | | | | | ||| | | ||| ||
Db     465 CATTGCCCGCTGGGACCCGCTGCTTGGCGACCACGGCGTGATCCAAGGGGACACGGCCCTAA  524

Qy     269 TCACCCGACCCGGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGACATGATGAAAGAAGCATTAG     328
           |   | |      |   || |||  |  | | | || || | || |   | | |||  ||
Db     525 TCACGCGGCCCTGAGCCCGCTCGTAGGCCAGAGGTCGATGAGTCCGCAAAGCCCTGC      584

Qy     329 AAAAACTTCAGCTCAATATAGTAGAGAATGAAAGATGAAAAATGCAACTTTAGATGGCGGAG     388
           ||  |  | | || |  ||  | | |  |  | || | ||  | |  |  |||  || |||
Db     585 AAGACCTGGGCGCTTCCGAATTGTGGAAATAGGAGACGAGAACCGCACGCTGGATGGCACTG  644

Qy     389 ATGTTTTATTCACAGGCAGAGAATTTTTTGTGGCCTTCCAAAAGGACAAATCAACGAG     448
           |  | ||  |  ||| | || |  || | |||| | | |  ||| |  | ||||| |||
Db     645 ACGTTCTCTTCACCGGCCGGGGAGTTTTTCGTAGGCCTCTCCAAATGACCAATCACCGAG  704

Qy     449 GTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGCCAG     508
           |  | ||  || |  | | | | ||   ||| | |  |||| || || |||| ||| |||
Db     705 GAGCTGAGATCGTGGCGGACACGTTCCGGGACTTCGCCTGCTCCCACTGTGCCAGTCTCGG  764

Qy     509 ATGGGTTGCATTTGAAGAGTTTCTGCGACGCATGGCGCATGGCGCCTAACCTGATCGCAATTGGGT     568
           |  | |  |  ||||| | |  ||| | | || | | ||| |  |  |||| || | ||| | |||
Db     765 GTCCCTCCCACCTGCGCGTCTTCTGCGGCATGGGCGCCGCACTGTTGTGGCAGGCA  824

Qy     569 CTAGTGAATCTGCACAGAGAGGCCCTAAGAATCATGCAACAGATGAGTGACCACCGCTACG     628
           | ||| || |  | | | |  ||| || |||||||| | | |   | | ||||| |  |||
Db     825 GCAGCGACGCTGCCCAAAAGGCTGTCCGGGAACGGTGCCAATGGCCAGATCACCCCATATG  884

Qy     629 ACAAACTCACTGTGCCTGATGACATAGACAGACAAACTGTATATATCTAAAT-----ATCC     682
           | |  || | | |  |  ||| | | | |  | |||| | | |  | | ||||     ||||
Db     885 CCTCCCTGACCCTCCCAGATGACGCAGCTGCACCGGAGGTGGGGATCTGCTTCGTCCTTGGGTTGC  944

Qy     683 CCAACAAAGGGCACGTCTTGCTGCACCGAACCCGGAGAGTATCCAGAAAGTGCAAAGG     742
           || | | || | | | || |  |  |||  | |  | | | | | ||| | |||| |||
Db     945 CTGGTGTGCCCCCTTTCCTTCCTGCACCGTGGAGGTGTGGGGGATCTGCCCAACACGCCAGGAGG  1004

Qy     743 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGG     802
           |  | || |  | |  | |   ||| | | |  || | | || | | |  || |  ||||
Db    1005 CACTGCAGAAGCTCTCTGATGTCATCACCCGGTACCGTGTCCTGCTCAGAACTGGAGAAGG  1064

Qy     803 TGGATGGGCTGCTCACCTGGCTCCTCAGTTTTAATTAACA     841
           |  | | ||| | | ||
Db    1065 CCGGCGCCGGGCTTCAGCTCCCCTGCTTGGTGCTCAGCA  1103
```

Search completed: June 7, 2004, 20:32:46
Job time : 412 secs

```
              Copyright  GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:       June  7, 2004, 19:08:46 ; Search time 2569 seconds
                                        (without alignments)
                                        9973.434 Million cell updates/sec

Title:        US-09-889-733B-1
Perfect score: 858
Sequence:     1 atggccggcctcggccaccc.........acaagaaggtagactcctga 858

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:     2751289 seqs, 1931090276 residues

Total number of hits satisfying chosen parameters:    55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    EST:*
         1:   em_estba:*
         2:   em_esthum:*
         3:   em_estin:*
         4:   em_estmu:*
         5:   em_estov:*
         6:   em_estpl:*
         7:   em_estro:*
         8:   em_htc:*
         9:   gb_est1:*
        10:   gb_est2:*
        11:   gb_htc:*
        12:   gb_est3:*
        13:   gb_est4:*
        14:   gb_est5:*
        15:   em_estfun:*
        16:   em_eston:*
        17:   em_gss_hum:*
        18:   em_gss_inv:*
        19:   em_gss_pln:*
        20:   em_gss_vrt:*
        21:   em_gss_fun:*
        22:   em_gss_mam:*
        23:   em_gss_mus:*
        24:   em_gss_pro:*
        25:   em_gss_rod:*
        26:   em_gss_phg:*
        27:   em_gss_vrl:*
        28:   gb_gss1:*
        29:   gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

```
                              SUMMARIES

Result          %      Query
No.    Score  Match  Length DB Length ID        Description
----   -----  -----  ------ -- ------ --         -----------
  1   717.2   83.6   1109   11 AK010430    AK010430 Mus muscu
  2   644.4   75.1    889   12 BI763990    BI763990 603049879
  3   631.4   73.6   1001   13 BY709864    BY709864 BY709864
  4   582.6   67.9    746   14 CB249709    CB249709 UI-M-EXO-
```

```
  5   577.8   67.3    862   13          BQ899468  BQ899468 AGENCOURT
  6   547.8   63.8    760   14          CA512485  CA512485 UI-R-FJO-
  7   544.4   63.8    822   9           AI647789  AI647789 uk43b05.x
  8   534.2   62.3    656   13          BY742062  BY742062 BY742062
  9   517.8   60.3    522   9           AI751463  AI751463 cn10b04.x
 10   480.8   56.0    698   13          BY754930  BY754930 BY754930
 11     479   55.8    731   10          BB611533  BB611533 BB611533
 12   472.2   55.0    800   9           AA986117  AA986117 uc81h06.x
 13   453.4   52.8    455   29          AY406329  AY406329 Homo sapi
 14   450.2   52.5    455   29          AY406330  AY406330 Pan trogl
 15   445.8   52.0    685   14          CB247154  CB247154 UI-M-FI0-
 16     439   51.2    702   13          BY731938  BY731938 BY731938
 17   430.6   50.2    645   29          CG477727  CG477727 OST7734 M
 18     428   49.6    654   9           AU296050  AU296050 AU296050
 19   425.4   49.6    624   10          AW514143  AW514143 hd24a02.x
 20   413.2   48.2   1035   10          BF579996  BF579996 602095987
 21   411.6   48.0    501   29          CG617513  CG617513 OST11187
 22   410.4   47.8    923   12          BI332932  BI332932 602982027
 23     396   46.2    777   14          CD349526  CD349526 UI-M-FY0-
 24   390.4   45.5    860   14          CF548604  CF548604 AGENCOURT
 25     388   45.2    486   9           AI463821  AI463821 va31a06.y
 26   386.4   45.0    806   9           AU051351  AU051351 AU051351
 27   381.8   44.5    681   12          BM963225  BM963225 UI-M-EQ0-
 28   380.4   44.3    569   9           AI265716  AI265716 uj05e03.x
 29   377.2   44.0    776   14          CD751679  CD751679 AGENCOURT
 30     375   43.7    455   12          AY406331  AY406331 Mus muscu
 31     375   43.7    556   12          BM964027  BM964027 UI-M-EQ0-
 32   370.8   43.2    638   13          BY713422  BY713422 BY713422
 33   361.6   42.1    555   9           AI050531  AI050531 uc83d09.y
 34   361.6   42.1    817   12          BI221174  BI221174 602939511
 35     359   41.8    792   14          CD750695  CD750695 AGENCOURT
 36   357.6   41.7    547   9           AI661738  AI661738 va31a06.x
 37   357.6   41.7    897   14          CK019367  CK019367 AGENCOURT
 38     357   41.6    432   10          BB864630  BB864630 UI-M-BH1-
 39     356   41.5    550   9           AA986603  AA986603 uc81h06.y
 40   352.8   41.1    356   12          BG977444  BG977444 RC5-CI016
 41   344.6   40.2    571   14          CB522710  CB522710 UI-M-GK0-
 42   344.6   40.2    571   14          CB522972  CB522972 UI-M-GK0-
 43   338.6   39.5    539   9           AW012296  AW012296 um07b05.y
 44   338.6   39.5    720   9           AV167374  AV167374 AV167374
 45   338.2   39.4    802   14          CA588435  CA588435 hab63f06.
```

                              ALIGNMENTS

```
RESULT 1
AK010430
LOCUS       AK010430              1109 bp    mRNA    linear   HTC 20-SEP-2003
DEFINITION  Mus musculus ES cells cDNA, RIKEN full-length enriched library,
            clone:2410006N07 product:dimethylarginine dimethylaminohydrolase 1,
            full insert sequence.
ACCESSION   AK010430
VERSION     AK010430.1  GI:12845867
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279953
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
```

REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,Y., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4 (bases 1 to 1109)
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 1109)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group; RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT   Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAGAGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTAATTAAATTAAATATCCCCCCCCCCCCCC 3']. cDNA was cleaved
            with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
            Host: SOLR.
FEATURES   Location/Qualifiers
     source   1..1109
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2410006N07"
                     /db_xref="MGI:1901307"
                     /db_xref="taxon:10090"
                     /clone="2410006N07"
                     /cell_type="ES cells"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
     CDS   137..994
                     /note="unnamed protein product; dimethylarginine
                     dimethylaminohydrolase 1 (MGD|MGI:1916469, GB|AK010430,

evidence: BLASTN, 100%, match=1109)
putative"
                     /codon_start=1
                     /protein_id="BAB26932.1"
                     /db_xref="GI:12845868"
                     /translation="MAGLGHPSAFGRATHAVVRAPPESLCRHALRRSQGEVDFARAE
                     RQHELYVGVLGSKLGLGQVVQLPADESLPDCVPVEDVAVVCETALITRPGAPSRRKEV
                     DMMKEALEKLQLNIVEMKDENATLDGGDVLFTGREFFVGLSKRTNQRGAEIIADTFKD
                     YAVSTVPVADSLHLKSFCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDM
                     AANCIYLNIPSKGHVLLHRTPEYPESAKVYEKLKDHLLIFVSNSEMEKVDGLLTCCS
                     VFINKKIDS"

ORIGIN

Query Match   83.6%; Score 717.2; DB 11; Length 1109;
Best Local Similarity   89.7%; Pred. No. 1.3e-123;
Matches 770; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy   1 ATGGCCGGCCTCGGCCACCCCTCGGCCGGCCACCCACCGCCGTGGTGCGGCG 60
          |||||||||||||||||||||||||||||||| || ||||||||||||||||
Db   137 ATGGCCGGCCTCGGCCACCCCTCGGCCGGCCACCCACCGCCGTGGTGCGGGCT 196

Qy   61 CTACCCGAGTCGCTCTGCCAGCACGGCGCGTCGAGAAGGCGCAAGGGCGAGGCAGGTGACGTC 120
          ||||||| |||| |||||||||||||| || ||||||||||||||| |||||||||||||||||
Db   197 CCGCCCGAGTCCCTGTGCCGCACGGCGCGTCGAGGGCGCTGCCAGGGCAGGAGCGAGGTGATTTC 256

Qy   121 GCCCGCGGCGAAGGGCCAGCACCAGCTCTACGTGGGCGTTGCTGGCGCAAGCAGTGGGGGGTG 180
          |||||||||||||| | || |||| |||| |||||||||||||||||||| || |||||||||
Db   257 GCTCGCGCCGAGCGCCAGCAGGAGCTCTACGTGGGCGTTGCTGGCGCAAGCAGTGGGGGGTG 316

Qy   181 CAGGTGGTGGAGCTGCCGGACGAGAGCCTTCGGACGAGCGTCTTCGTGGAGGACGTG 240
          ||||||||||||||| |||||||||||||||||||||||||||||||| ||||||||||
Db   317 CAGGTGGTGGCAGCTGCCGCCGACGAGAGCCTGCCGGACTGCGTGCCGGTTGCTGGAGGACGTG 376

Qy   241 GCCGTGGTGCGGAGGAGACGGGCCCTCATCACCCGACCCCGGGCGGCCGAGCCGGAGAAG 300
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   377 GCCGTCGTGCGGAGGAGACGGGCCCTCATCATCACCCGCCCCGGGCGGCCCAGCCCGGAGAAG 436

Qy   301 GAGGTTGACATGATGAAGGCATTAGAAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||| |||
Db   437 GAGGTTGACATGATGAAGGCATTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAA 496

Qy   361 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420
          |||||||||||||||||||||||||||||||||| ||||||||||||| ||||||||||
Db   497 GATGAAAATGCAACTTGATGGGGACGTCCTATTCACAGGCAGAGAATTTTTTGTG 556

Qy   421 GGCCTTTCCAAAAGGACAAAATCAACGAGGTGCTGAAATCTTGGCTGATTACTTTTAAGGAC 480
          |||||||||||||||||||| |||||||||||||||||||||||||||||||| |||||||||
Db   557 GGCCTTTCCAAAAGACAAAATCAAACGAGGTGCTGAAATTCTGGCTGATTACTTTTAAGGAC 616

Qy   481 TATGAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATG 540
          |||||| ||||||||||||||| |||||||||||||| |||||||||||||||||| |||
Db   617 TACGAGTCTCTACAGTGCCGATTCCTGTGGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCATG 676

Qy   541 GCTGGGCCTAACCTGATCGCAATTGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   677 GCCGGACCCAACCTGATTGCAATAGGTCCAGGACAACTGCACAGAAGGCCCATCAAGATC 736

Qy   601 ATGCAACAGATGAGTGAGTGACCACCATCGTTATGACAAGCTCACTGTACCCGACACATAGCAGCA 660
          |||||| || || |||||||| ||| ||||| |||| ||||||||||||||||||||||||| |||
Db   737 ATGCAACAGATGAGTGAGTGACCATCGTTATGACAAGCTCACTGTACCCGACACATGCCGCC 796

Qy   661 AACTTGTATATATCTAAATATCCCCAACAAAGGACCAGTCTTGCTGCCACCGAACCCCGGAA 720
          ||||| |||||||| || |||||||||||||||| ||| ||||| |||||||||||||| |||
Db   797 AACTTGCATATATCTAAATATCCCCAGCAAAGGGCATGTCTTGTCTGCCACCGAACCCCAGAA 856

Qy   721 GAGTATCCAGAAGTGCAAAGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTG 780
          ||||| |||||| ||||||| ||| |||||||||| |||||||||||||||||||||||
Db   857 GAGTACCCAGAAGCGCAAAGGTCTATGAGAAACTCAAGGACCATCTACTACTCCTGTG 916

Qy   781 AGCATGTCTGAACTGGAAAAGTGGATGGGCTCGCTCACCTGCTGCTCAGTTTTAATTAAC 840
          ||| ||| ||||||||||||||| || |||||||||||||| ||||| ||||||||||||||
Db   917 AGCAACTCGGAGTGGAAAAGGTGGACGGCTTGCTCACCTGCTGCTCCGTTTTTATTAAC 976

```
Qy  241 GCCGTGGTGTGCGAGGAGACGCGCCCTCATCACCCGGGGCGGCCGAGCCGGAGGAG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  342 GCCGTGGTGTGCGAGGAGACGCGCCCTCATCACCCGGGGCGGCCGAGCCGGAGGAG  401

Qy  301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTGAGATGAAA  360
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  402 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA  461

Qy  361 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG  420
            |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  462 GATGAAAATGCAACTTAGATGGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG  521

Qy  421 GGCCTTTCCAAAAGACAAATCACGAGTGCTGAAATCTTGGCTGGATACTTTAAGGAC  480
            |||||||||||||| |||||||||| ||||||||||||||||||||||||||||||
Db  522 GGCCTTTCCAAAAGGACAAATCACCAGGGTCTGAAATCTTGGCTGGATACTTTAAGGAC  581

Qy  481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAGAGAG-TTTCTGCAGCAT  539
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  582 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAGAGAGTTTTCTGCAGCAT  641

Qy  540 GGCTGGGCCTAACCTGATCGGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGAT  599
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  642 GGCTGGGCCTAACCTGATCGGCAAATGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGAT  701

Qy  600 CATGGCAACAGA-TGAGTGACCACCCTACGAC-AAACTCACTGT-GCCTGATGACATAGC  656
            |||| ||||||| ||||||||| ||| |||||  |||||||||| |||||||||||||
Db  702 CATGCAACAGATTGAGTGACGAACCCTACGACGACAAAACTCACTGTGGCCTGATGACGTTAG  761

Qy  657 AGCAAACTGTATATATCT-----AAATATCCCCACAAAGGGACACGTCTTGCTGCCACCG  710
            ||| ||| ||||||||||     ||| |||||||||||||||||||||||||||||||||
Db  762 CAGCAAACGTGTATATATTCTAAATTATTCCCCACAAAGGGGCCCATCTGTGGCCTTGGACG  821

Qy  711 AACCCCGGAAGAGTATCCAGAAAGTGCAAAGTTTATGGACAAAACTGAAGAACCATATGCT  770
            ||||||||||||||||||||||||||||||| ||||| ||||| |||| |||| ||| |
Db  822 GAACCCGGAAGAGTATCCAGAAAGTGCAAAGGTTATGACAAACTTGAGGACCTAATGGT  881

Qy  771 GATCCCCG  778
            ||||||||
Db  882 GATCCCCG  889
```

```
Qy  841 AAGAAGGTAGACTCCTGA  858
            ||||| ||||||||||||
Db  977 AAGAAGATAGACTCCTGA  994
```

Query Match 75.1%; Score 644.4; DB 12; Length 889;
Best Local Similarity 92.9%; Pred. No. 4.7e-110;
Matches 732; Conservative 0; Mismatches 46; Indels 10; Gaps 5;

```
Qy    1 ATGGCCGGCCGCTCGGCCACCCCCTCGGCCGTTCCGGCCGGGCCGCACCCACCGCCGTGGTGCCGGGCG  60
            ||||||||| |||||||||||||  |||| |||||||||||||||||||||||||||||||||||||
Db  102 ATGGCCGGGCTCGGCCACCCCCGCCGCCGCTTCGGCCGCCGCACCCACCGCCGTGGTGCCGGGCG  161

Qy   61 CTACCCGAGTCGCTCTGCCAGCACGCGCACGCGCGGGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC  120
            |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  162 CTACCCGAGTCGCTCGGCCGCACGCGCACGCGCGGGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC  221

Qy  121 GCCCGCGGGACGGCAGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGTCGGGGCTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  222 GCCCGCGGGACGGCAGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGTCGGGGCTG  281

Qy  181 CAGGTGGTGGAGCTGCCCGGCCGACGAGAGCGCCTTCCGGGACTGCGTCTTCGTGGAGGACGTG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  282 CAGGTGGTGGAGCTGCCCGGCCGACGAGAGCGCCTTCCGGGACTGCGTCTTCGTGGAGGACGTG  341
```

Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,L., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE   Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs

JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
          Laboratory for Genome Exploration Research Group, RIKEN Genomic
          Sciences Center (GSC), Yokohama Institute
          The Institute of Physical and Chemical Research (RIKEN)
          1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          Tel: 81-45-503-9222
          Fax: 81-45-503-9216
          Email: genome-res@gsc.riken.go.jp,
          URL:http://genome.gsc.riken.go.jp/
          Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
          Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
          Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
          Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
          Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
          Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
          Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
          Direct Submission
          Computational Analysis of Full-Length Mouse cDNAs Compared with
          Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
          Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new
          genes. Genome Res. 10 (10), 1617-1630 (2000)
          RIKEN integrated sequence analysis (RISA) system-384-format
          sequencing pipeline with 384 multicapillary sequencer. Genome Res.
          10 (11), 1757-1771 (2000)
          Computer-based methods for the mouse full-length cDNA
          encyclopedia: real-time sequence clustering for construction of a
          nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
          cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Please visit our web site (http://genome.gsc.riken.go.jp) for
          further details.
FEATURES   Location/Qualifiers
     source   1..1001
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="2410006N07"
                     /cell_type="ES cells"
                     /lab_host="SOLR"
                     /clone_lib="RIKEN full-length enriched, ES cells"
                     /note="Site_1: XhoI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAGAGAGAGACGTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 5.0 and subtraction to Rot = 25.0. Second strand
                     cDNA was prepared with the primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAATTAATTAATCCCCCCCCCCCC 3']."

ORIGIN

```
Qy          1 ATGGCCGGCTCGGCCACCCCTCGCCTTCGGCGGCGGCCACCCACGCCGCCGTGGTGCGGCCG 60
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        134 ATGGCCGGCTCGGCCACCCCTCGCCTTCGGCGGCGGCCACCCACGCCGCCGTGGTGCGGGCT 193

Qy         61 CTACCCAGTCGTCTGCCAGCACGCGTCGAGAAGCGCCAAGGCGCGAGGAGGTGGACGTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||| | ||||||||
Db        194 CCGCCCGAGTCCCTGTGCCCACCGCGTCGAGGCGCCTCGCAGGGCGAGGAGGTGGATTTC 253

Qy        121 GCCCGCGGGAACGGCCAGCACC-AGCTTACG-TGGGGTGCTGGGCAGCAAGCTGGGGC 178
              |||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||
Db        254 GCTCGCGCCGAGCGCCAGCACCGGAGCTCTACGTCGGGGTGTCGTGGGCAGCAAGCTGGGGC 313

Qy        179 TGCAGGTGGTGGAGCTGCCGGACGAGAGCCTTCGGACGCGTCTTCGTCGTGGAGGACG 238
              ||||||||||||||||||||||||||||||||||||||| | |||||||||||||||||
Db        314 TGCAGGTGGTGCAGCTGCCCGGACGAGAGCCCTGCCGACCTCGTGTCGTCGTGGAGGACG 373

Qy        239 TGGCCGTGGTGTGCGAGGAGACGGCCCCATCACCCGGGGCCCCGGAGCCGGAGGA 298
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        374 TGGCCGTGTGTGCGAGGAGACGGCCCCTCATCACCACCCGCCCCAGCCCGCAGGA 433

Qy        299 AGGAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGA 358
              ||||||||||||||||||||||| |||||||||||||||||||||||| ||||||||||||
Db        434 AGGAGGTTGACATGATGAAAGAAGAGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGA 493

Qy        359 AAGATGAAAATGCAACTTTAGATGGCGAGATGTTTATTCACAGGCAGAGAATTTTTG 418
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        494 AAGATGAAAATGCAACTTTGGATGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTG 553

Qy        419 TGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAGG 478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 TGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAGG 613

Qy        479 ACTATGCAGTCTCCACAGTGCCAGTGCCAGATGGGTTGCATTGAAGAGTTTCTGCAGCA 538
              ||||||||||||||||||||||||| ||||| |||||||||||||||||||||||||||
Db        614 ACTACGCAGTCTCACAGTCGCGTGGGCCGATTCTTCGCATTTAAAGAGTTTCTGCAGCA 673

Qy        539 TGGGTGGGCTAACCTGATCGCAATTGGGTCTAGTAGAATCGACAGAGAGGCCCTTAAGA 598
              | |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        674 TGGCCGGACCCAACCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAGAGGGCTCAAGA 733

Qy        599 TCATGCAACGATGAGTGACCACCGCTACGACAAACCTCACTGTGCCTGATGACATAGCAG 658
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        734 TCATGCAACAGATGAGTGACCATCGTTATGACAGCAGCTCACTGTACCCGACGACATGGCCG 793

Qy        659 CAAACTGTATATATCTAAATATCCCACACAAAGGGCACGTCTTGCTCACCGAACCCCGG 718
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        794 CCAACTGCATATATCTAAATATCTAAATATCCCCAGCATAGGGCATGTCTTGCTCACTGAACCCCAG 853

Qy        719 AAGAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACCATATGTGATCCCCG 778
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        854 AAGAGTACCCAGAAAGCGCAAACGTCTATGAGAACTCAAGGACCATCTCAAGGACCATCTCG 913

Qy        779 TGAGCA 784
              ||||||
Db        914 TGAGCA 919
```

RESULT 4
CB249709
LOCUS   CB249709   746 bp mRNA   linear EST 15-JUL-2003
DEFINITION UI-M-EXO-byl-1-07-0-UI.r1 NIH_BMAP_EXO Mus musculus cDNA clone
            IMAGE:5719830 5', mRNA sequence.
ACCESSION   CB249709
VERSION   CB249709.1 GI:28389411
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus

REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE   1 (bases 1 to 746)
JOURNAL   NIH-MGC http://mgc.nci.nih.gov/.
COMMENT   National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Dr. James Lin, Univeristy of Iowa
              cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
              cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: MGC clone distribution infomation can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              This clone was contributed by the Brain Molecular Anatomy Project
              (BMAP)

FEATURES   Seq primer: pYX-5.
source   Location/Qualifiers
              1..746
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6"
              /db_xref="taxon:10090"
              /clone="IMAGE:5719830"
              /tissue_type="whole brain"
              /dev_stage="embryo 15.5 dpc"
              /lab_host="DH10B (T1 phage resistant)"
              /clone_lib="NIH_BMAP_EXO"
              /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
              Site_2: Not I; The library was constructed according to
              Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
              1996. Denatured mRNA was size fractionated on a 1% agarose
              gel. First strand cDNA synthesis was primed with an
              oligo-dT primer containing a Not I site. Double stranded
              cDNA was size selected according to mRNA size fraction,
              ligated with EcoR I adaptor, digested with Not I, and then
              cloned directionally into pYX-Asc vector. The library tag
              sequence located between the Not I site and the polyA
              tail, is GTGCGTGGAA. This library was created for the
              University of Iowa Mouse Brain Molecular Anatomy Project
              (BMAP) 'Gene Discovery in the Developing Mouse Nervous
              System', supported by National Instutites of Mental Health
              (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match   67.9%;   Score 582.6;   DB 14;   Length 746;
Best Local Similarity   89.8%;   Pred. No. 1.5e-98;
Matches 624;   Conservative   0;   Mismatches 71;   Indels   0;   Gaps   0;

Qy   164 GCAGCAAGCTGGGCTGCAGGTTGGTGAGCTGCCGGCCGACGAGAGGCCTTCCGGACTGCG 223
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   1 GCAGCAAGCTGGGGCTGCAGGTTGGTGAGCTGCCGCCCGCCGACGAGAGCCTGCCCGACTGCG 60

Qy   224 TCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGGACGGCCCTCATCACCCGGACCCGGGG 283
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGTTCGTGGAGGACGTGGCCGTGTGTGCGAGGAGGACGGCCCTCATCACCCGCCCCGGGG 120

Qy   284 CGGCCGAGCCGGAGGAAGGAGGTTGACATGAGGAGCATTAGAAAAGAGCATTAGAAAAAACTTCAGCTCA 343
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CGCCCAGCCGCAGGAGGAGGTTGACATGATGAAAGAGCATTAGAAAAACTTCAGCTCA 180

Qy   344 ATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATGCCGGAGATGTTTATTCACAG 403
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACATAGTAGAGATGAAAGATGAAAATGCAACTTTGGATGGTGGGGACGTCCTATTCACAG 240

Qy   404 GCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGTGCTGAAATCTTGG 463
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GCAGAGAATTTTTGTGGGCCTTTCCAAAAGAACAAATCAACGAGTGCTGAAATCTTGG 300

Qy   464 CTGATACTTTAAGGACTATGCAGTCCCACAGTGCCAGTGCAGATGGGTTGCATTGA 523
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTGATACTTTAAGGACTACGCAGTCTCTACAGTGCCCTGTGCGAGTTCTTTTTGCATTTAA 360

REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE   1 (bases 1 to 746)
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   NIH-MGC http://mgc.nci.nih.gov/.
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Dr. James Lin, Univeristy of Iowa
              cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
              cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Distribution: MGC clone distribution infomation can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              This clone was contributed by the Brain Molecular Anatomy Project
              (BMAP)

Qy   524 AGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCAC 583
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGAGTTTCTGCAGCATGGCTGGGCCTCAACCTGATTGCTCAATAGGGTCCAGCGAATCTGCAC 420

Qy   584 AGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGC 643
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGAAGGCCCTCAAGATCATGCAACAGATGAGTGACCATCGGTATGACAAGTCACTGTAC 480

Qy   644 CTGATGACATAGCAGCAAACTGTATATATCCAAATATCCCAACAAAGGGCACGTCTTGC 703
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CCGACGACATGGCCCGCCAACTGCATATATTAAATCCCAGCANAGGGCATGTCTTGC 540

Qy   704 TGCACCGAACCCGGAAGAGTATCCAGAAGTGCAAAGGTTTATGAGAAACTGAAGGACC 763
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TGCACCGAACCCCAGAGAGTACCCAGAGAAGCGCANAGGTCTATGAGAAACTCAAGGACC 600

Qy   764 ATATGCTGATCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTTGCTCACCTGCT 823
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATCTACTGATCCCTGTGAGCAACTCGGAGATGAAAAGGTGGACGGCTTGCTCACCTGCT 660

Qy   824 GCTCAGTTTAATTAACAAGAGAGGTAGACTCCTGA 858
         |||||||||||||||||||||||||||||||||||
Db   661 GCTCCGTTTTTATTAACAAGAAGATAGACTCCTGA 695

RESULT 5
BQ899468
LOCUS   BQ899468   862 bp   mRNA   linear   EST 16-AUG-2002
DEFINITION   AGENCOURT_8750357 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6332815
              5', mRNA sequence.
ACCESSION   BQ899468
VERSION   BQ899468.1   GI:22291482
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 862)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
              Ph.D.
              cDNA Library Preparation: ResGen, Invitrogen Corp
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM13791 row: a column: 08
              High quality sequence stop: 704.

FEATURES   Location/Qualifiers
source   1..862
              /organism="Mus musculus"
              /mol_type="mRNA"
              /db_xref="taxon:10090"
              /clone="IMAGE:6332815"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_130"
              /note="Organ: otocysts; Vector: pCMV-SPORT6.1; Site_1:
              EcoRV, Site_2: NotI; Cloned unidirectionally; Primer:
              Oligo dT. Average insert size 1.95 kb. Constructed by
              ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN

Query Match   67.3%;   Score 577.8;   DB 13;   Length 862;
Best Local Similarity   89.6%;   Pred. No. 1.3e-97;
Matches 621;   Conservative   0;   Mismatches 72;   Indels   0;   Gaps   0;

Qy   1 ATGGGCGGCCTCGGCCACCCTCCGCCTTCGGCCGCGGCCACCCACGCCGCGGTGGTGCGGGGG 60

```
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. James Lin, Universtiy of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
    Genetics (www.resgen.com).
Seq primer: M13 REVERSE.

FEATURES             Location/Qualifiers
     source          1..760
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-FJ0-cpz-1-06-0-UI"
                     /tissue_type="embryo"
                     /dev_stage="embryo"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-R-FJ0"
                     /note="Vector: pYx-Asc; Site_1: EcoR I; Site_2: Not I;
                     UI-R-FJ0 is a cDNA library containing the following
                     tissue(s): rat embryo. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dt)18 tail. The
                     sequence tag for this library is CATCTCACT. This library
                     was created for the University of Iowa Program for Rat
                     Gene Discovery and Mapping (Val Sheffield, Bento Soares
                     and Tom Casavant)"

ORIGIN
Query Match             63.8%;  Score 547.8;  DB 14;  Length 760;
Best Local Similarity   89.8%;  Pred. No. 4.8e-92;
Matches   588;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy   204 CGAGAGCCTTCCGACTCGTCTTCGTGGAGACGTGCCGTGTGTGCGAGGAGACGGC 263
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     1 CGAGAGCCTTCCGACTCGTCTTCGTGGAGACGTGCCGTGTGTGCGAGGAGACCGC 60

Qy   264 CCTCATCACCGACCCGGACGCCGGCCCGGAGGAAGGAGGTTGACATGATGAAAGAGAGC 323
             ||||||||||| |||||||||||  || ||||||||||||||||||||||||||||||
Db    61 CCTCATCACCGCCCCGGACGCCGGGGCGGCCGCGGAGGAAGGAGGTTGACATGATGAAAGAGAGC 120

Qy   324 ATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAGATGAAAATGCAACTTTAGATGG 383
             |||||||||||||||||  |||||||||||||||||||||||||||||| ||||||||
Db   121 TTTGAAAAACTTCAGCTCAACATAGTAGAGATGAAGATGAAAATGCAACCTTAGATGG 180

Qy   384 CGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCAAAAGGACAAATCA 443
             |||||| |||||||||||||||||| ||||||||||||||||||||||||||||||
Db   181 TGGGGACGTCCATTCACAGGCAGAGAGTTTTTTGTGGGCCTTTCAAAAGGACAAATCA 240

Qy   444 ACGGAGGTGTGAAATCTTGGCTGATACTTTTAAGGACCATGCAGTCTCCACAGTGCCAGT 503
             |||||||||||| |||||||||||||||||  ||||||  |||||  |||||| | |||
Db   241 ACGGAGGTGTGAGATCTTGGCTGATACTTTCAAGGACTACACGCAGTTTCCACAGTCCCCGT 300

Qy   504 GGCAGATGGGTTGCATTTGAAGAGTTTCGCAGCAGTTTCTGCAGGCATGCTGGGCCTAACCTGATCGCAAT 563
             ||| |||  || |||||| |||||||||  |||| |  |||||  ||| || ||||| ||||||||||||
Db   301 GGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCATGGCCATGGCTGGGCCCAACCTGATCGCAAT 360

Qy   564 TGGGTCTCAGTGAATCTGCACAGAAGGCCCTTAAGATCATCGACAGATGAGTGACCACCG 623
             ||||||   |||||||| ||||||||||||||  ||| || |||||||||| |||||||
Db   361 AGGGTCCAGTGAATCTGCCAGATCATCGCCCTCAAGATCATGACAGATGAGTGACCACCG 420

Qy   624 CTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCC 683
             ||| |||||||||||||||||||| ||||  |||||  |||||||||||||||||||
Db   421 TTATGACAAGCTCACTGTACCGGACGCATGGCCGCCAACTGGCCGCCAACTGTATATATTTAAATATCCC 480
```

```
Db   131 ATGGCCGGCCCTCCGGCCACCCCTCCGCCTTCGCCCGGACCACCCAGCCCGGGTGCGGGCT 190
             ||||||||||||||||
Qy    61 CTACCCGAGTCGCTGCCAGCACGCGGCTGCCAAGGGGCGCCAAGGGCCGAGGAGGTGGACGTC 120
             |||||||||||
Db   191 CCGCCCAGTCCCTGTGCCGCCACGCGGCTGAGGCGCGCTGAGGCGGCGGCTGGATTTC 250
             |||||||||
Qy   121 GCCCGCGGGAACGGCAGCAGCAGCAGCACCAGTCTCACGTGGCGTGGGCAAGCTCGGGGCTG 180
             |||||||||||
Db   251 GCTCCGGCCGAGGCGGCACGCAGCTCACGTGCGGCGTGCTGGGCCAGCAGCAAGCTGGGGCTG 310
             ||||||||||||
Qy   181 CAGGTGTGGAGTGCGGCGGCCGACGAGCCTTCCGGACTGCGTCTTCGTGGAGACGTG 240
             |||||||||||
Db   311 CAGGTGTGGAGTGCGGCGGCCGACGAGCCTGCCGGACTGCGTCTTCGTGTCGTGGAGACGTG 370
             ||||||||||||
Qy   241 GCCGTGTGTGCGAGGAGACGGCCCTCATCACCGACCCGGACGCCGGGGCGGCCGAGGAAG 300
             |||||||||||
Db   371 GCCGTGTGTGCGAGGAGACGGCCCTCATCACCGCCCCCGGACGCCGGGGCGGCCGCGGAGGAAG 430
             ||||||||||||
Qy   301 GAGGTTGACATGATGAAAGAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
             |||||||||||||
Db   431 GAGGTTGACATGATGAAAGAGAGCTTTGAAAAAACTTCAGCTCAACATAGTAGAGATGAAA 490
             ||||||||||||||
Qy   361 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420
             |||||||||||
Db   491 GATGAAAATGCAACTTTGGATGGTGGGGACGTCCATTCACAGGCAGAGAGTTTTTGTG 550
             ||||||||||||
Qy   421 GGCCTTTCAAAAGGACAAATCAACGGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
             |||||||||||
Db   551 GGCCTTTCAAAAGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 610
             ||||||||||||
Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCGCAGCATG 540
             |||||||||||
Db   611 TACGCAGTCTCTCACAGTCCCGTGGCCGATTCTTTGCATTTAAAGAGTTTCGCAGCATG 670
             ||||||||||||
Qy   541 GCTGGGCCTAACCTGATCGCAATTGGTCTCAGTGAATCTGCACAGAAGGCCCTTAAGATC 600
             |||||||||||
Db   671 GCCGGACCCAACCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAAGGCCCTCAAGATC 730
             ||||||||||||
Qy   601 ATGCAACAGATGAGTGAGTGACCACCGCTACGACGACAACTCACTGTGCCTGATGACATAGCAGCA 660
             |||||||||||
Db   731 ATGCAACAGATGAGTGAGTGACCATCGTTATGACAAGCTCACTGTATATATCTAAATATCCCCACAAAGG 790
             |||||||||||
Qy   661 AACTGTATATATCTAAATATCCCCACAAAGG 693
             ||||||||
Db   791 CACTGCATATATCTAAATATCCCCAGCAAAGG 823
```

RESULT 6
```
CA512485
LOCUS       CA512485                 760 bp    mRNA    linear   EST 15-NOV-2002
DEFINITION  UI-R-FJ0-cpz-1-06-0-UI.r1 UI-R-FJ0 Rattus norvegicus cDNA clone
            UI-R-FJ0-cpz-1-06-0-UI 5', mRNA sequence.
ACCESSION   CA512485
VERSION     CA512485.1  GI:25003439
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
```

```
Query Match      63.4%;   Score 544.4;   DB 9;    Length 822;
Best Local Similarity   88.2%;   Pred. No. 2.1e-91;
Matches  623;   Conservative   0;   Mismatches  80;   Indels   3;   Gaps   3;

Qy     154  GGCGTGCTGGCAAGCTGGGGTGCAGTGGTGCAGTGGTGGAGCTGCCGGCCGACGAGAGCCTT  213
            ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     822  GGCGTTGCTGGCAAGCTGGGGNCTGCAGTGGTGCAGTGGTGCAGTGGTCCCCGACGAGAGCTG  763

Qy     214  CCGGACTGGCGTTCGTGGAGGACGACGTGGTGTGCGAGGACGAGGCGGCCTCATCACC  273
            |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     762  CCGGACTGGCGTTCGTGGAGGACGACGTGGTGTGCGGA-GAGACGGGCCTTCATCACC  704

Qy     274  CGACCC-GGGGCGCCCGAGCCCGGAGGAGGAGGTTGACATGATGAAAGAAGCATTAGAAAA  332
            |||||  ||| ||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     703  CGCCCCGGGGGCGCCCAGCCNGCAGGAAGGAGGTTGACATGATGAAAGAAGAGAGC-TTTGGAAA  645

Qy     333  ACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATGCGGAGATGT  392
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     644  ACTTCAGCTCAACATAGTAGAGATGANAGATGAAAAATGCAACTTTGGATGTGTGGGAGCGT  585

Qy     393  TTTATTCACAGGCAGAGAATTTTTTGTGGGCCTTTCCAAAAGGACAAATCAAGAGGTGC  452
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     584  CCTATTCACAGGCAGAGAATTTTNTNTGTGGGCCTTTCCAAAAGAACAAATCAACGAGGTGC  525

Qy     453  TGAAATCTTGGCTGATACTTTTAAGGACTACGCAGTGTCTCCACAGTGCCAGTGGCAGATGG  512
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db     524  TGAAATCTTGGCTGATACTTTTAAGGACTACTTTACAGTGCCCTGTGGCCGATTC  465

Qy     513  GTTGCATTGAAGAGTTCTGCAGATGGCTGGGCCTAACCTGCTGATCGCAATTGGTCTAG  572
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     464  TTTGCATTTAAAGAGTTCTGCAGATGGCTGGGCCGCACCCACCTGATTGCAATAGGGTCCAG  405

Qy     573  TGAATCTGCACGAAGGCCCCTTAAGATCATGCAACAGATGAGTGACCACCCTACGACAA  632
            |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     404  CGAATCTGCACGAAGGCCCTTCAAGATCATGCAACAGATGAGTGACCATCGTTATGACAA  345

Qy     633  ACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCCACAAAGG  692
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     344  GCTCACTGTACCGACGACATGGCCCCAACTGCCATATCGCATAATATCCCCAGCAAAGG  285

Qy     693  GCACGTCTTGCTGCACCGGACCCCAGACCCCAGAAGAGTATCCGAAGAGTGCAAAGGTTTATGAGAA  752
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     284  GCATGTCTTTGCTGCACCGGACCCCAGAAGAGTACCCAGAAGAAGCGCAAAGGTCTATGAGAA  225

Qy     753  ACTGAAGGACCATATGCTGATCCCCTGAGCATGTCGAACTGGAAAAAGTGGATGGGCT  812
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     224  ACTCAAGGACCATCATCACTCACTGATCCCCTGTGAGCAACTCCGGAGATGGAAAAGGTGGCACGGCTT  165

Qy     813  GCTCACCTGCTGCTCAGTTTTAATTAACAAGAAGTAGACTCCTGA  858
            |||||||||||||||||||||||||||||||||||||||||||||
Db     164  GCTCACCTGCTGCTCCGTTTTTATTAACAAGAAGATAGACTCCTGA  119
```

```
RESULT 8
BY742062
LOCUS       BY742062             656 bp    mRNA         linear       EST 17-DEC-2002
DEFINITION  BY742062 RIKEN full-length enriched, 2 days neonate sympathetic
            ganglion Mus musculus cDNA clone 7120491N07 5', mRNA sequence.
ACCESSION   BY742062
VERSION     BY742062.1  GI:27167348
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 656)
AUTHORS     Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogani,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
```

```
Qy     684  CAACAAAGGCACGTCTTGCTGCTGCACCGAACCCCGGAGAGTATCCAGAAAGTGCAAAGGT  743
            | |||||||||||||||||||||||| ||||||||||||||||||||| ||||||||||||
Db     481  CAGCAAAGGCATGTCTTGCTGCACCGAACCCCAGGAGAGTACCCAGAAAGCGCAAAGGT  540

Qy     744  TTATGAGAAACTGAAGGACCATATGCTGATCCCCTGAGCATGTCTGAACTGGAAAAAGGT  803
            ||||||||||||||||||||||||||| |||||||||||||||| |||||||||||||||
Db     541  TTATGAGAAGCTCAAGGACCATCATCACTCACTGATCCCCTGTGAGCAATTCTGAGATGGAAAAAGGT  600

Qy     804  GGATGGGCTGCTCACCTGCTGCTCAGTTTATTAATTAACAAGAAGGTAGACTCCTGA  858
            |||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     601  GGACGGCTTGCTCACCTGCTGCTCCGTTTTATTAACAAGAGACAGACTCTTGA  655

RESULT 7
AI647789/c
LOCUS       AI647789             822 bp    mRNA         linear       EST 30-APR-1999
DEFINITION  uk43b05.x1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1971729 3' similar to TR:O08557 O08557
            N-G,N-G-DIMETHYLARGININE DIMETHYLAMINOHYDROLASE. ;, mRNA sequence.
ACCESSION   AI647789
VERSION     AI647789.1  GI:4726467
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 822)
AUTHORS     Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
TITLE       The WashU-NCI Mouse EST Project 1999
JOURNAL     Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:988469
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end
            Possible reversed clone: similarity on wrong strand
            Seq primer: custom primer used
            High quality sequence stop: 516.
FEATURES             Location/Qualifiers
     source          1..822
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1971729"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse kidney mkia"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
                     (CACTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTTGGCCTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing; 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
ORIGIN
```

Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S.,Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Perea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE   Analysis of the mouse transcriptome based on functional annotation
           of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
           Laboratory for Genome Exploration Research Group, RIKEN Genomic
           Sciences Center(GSC), Yokohama Institute
           The Institute of Physical and Chemical Research (RIKEN)
           1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
           Tel: 81-45-503-9222
           Fax: 81-45-503-9216
           Email: genome-res@gsc.riken.go.jp,
           URL:http://genome.gsc.riken.go.jp/
           Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
           Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
           Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
           Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
           Nomura,K., Numazaki,R. Ohno,M., Ohsato,N., Saito,N., Sakazume,N.,
           Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
           Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
           Direct Submission
           Computational Analysis of Full-Length Mouse cDNAs Compared with
           Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
           Normalization and subtraction of cap-trapper-selected cDNAs to
           prepare full-length cDNA libraries for rapid discovery of new
           genes. Genome Res. 10 (10), 1617-1630 (2000)
           RIKEN integrated sequence analysis (RISA) system-384-format
           sequencing pipeline with 384 multicapillary sequencer. Genome Res.
           10 (11), 1757-1771 (2000)
           Computer-based methods for the mouse full-length cDNA
           encyclopedia: real-time sequence clustering for construction of a
           nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
           cDNA library was prepared and sequenced in Mouse Genome
           Encyclopedia Project of Genome Exploration Research Group in Riken
           Genomic Sciences Center and Genome Science Laboratory in Riken.
           Division of Experimental Animal Research in Riken contributed to
           prepare mouse tissues.
           RNA was provided by Dr. Akira Nakagawara (Div. of Biochemistry,
           Chiba Cancer Center Research Institute, 666-2 Nitona, Chuoh-ku,
           Chiba, 260-8717 Japan) whose assistance is gratefully acknowledged.
           Please visit our web site (http://genome.gsc.riken.go.jp) for
           further details.
FEATURES   Location/Qualifiers
     source   1..656
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="712049IN07"
                     /sex="mixed"
                     /tissue_type="sympathetic ganglion"
                     /dev_stage="2 days neonate"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 2 days neonate
                     sympathetic ganglion"
                     /note="Site 1: SalI; Site 2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAAGAAGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA
                     was cleaved with XhoI and BamHI Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC I.
                     Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
                     -RNA was provided by Akira Nakagawara, Div. of
                     Biochemistry, Chiba Cancer Center Research Institute,
                     666-2 Nitona, Chuoh-ku, Chiba, 260-8717 Japan, whose
                     assistance we gratefully acknowledge."

ORIGIN

Query Match   62.3%; Score 534.2; DB 13; Length 656;
Best Local Similarity   89.5%; Pred. No. 1.6e-89;
Matches   586; Conservative   0; Mismatches   68; Indels   1; Gaps   1;

```
Qy        182 AGTGGTGGAGCTGCCGGACGAGCAGAGCCTTCCGGACGAGCGTCTTCGTGGAGACGTGG 241
                ||||||| |||||||| ||||||||||||| |||||||||||||||||||||||||||||
Db          1 AGGTGGTGCAGCTGCCCGGACGAGCAGAGCCTTCCGGACGAGCGTCTTCGTGGAGACGTGG 60

Qy        242 CCGTCGTGTGCGGAGGAGACGGCCCTCATCACCCGGACCCGGGCGGGCCGGAGGAAGG 301
                |||||||||||||||||||||||||||||| |||| |||||||| |||||| ||||||
Db         61 CCGTCGTGTGCGGAGGAGACGGCCCTCATCACCCCGGGGGCCCGGGGCCCAGCCGGAGGG 120

Qy        302 AGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAG 361
                |||||||||||||||||||| |   |  ||||||||||||||| |||||||||||||||
Db        121 AGGTTGACATGATGAAAGAAGCGTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAG 180

Qy        362 ATGAAAATGCAACTTAGTAGTGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGG 421
                ||||||||||||||| | || ||| |||| || |||||||||||||||||||||||||||
Db        181 ATGAAAATGCAACTTGGATGGTCGGGAGACGTCCTATTCACAGGCAGAGAATTTTTGTGG 240

Qy        422 GCCTTCAAAAGACAAATCACGAGTGCTGAAATCTTGGCTGATACTTTTAAGGACT 481
                ||||| ||||||||||||| |||| ||||||||||||| ||||||||||||||||
Db        241 GCCTTTCAAAAGACAAATCAACGAGGTGCTGAAATCTTGGCGTGATACTTTTAAGGACT 300

Qy        482 ATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTCTCGCAGCATGG 541
                | |||||||| |||||  |  ||||| ||| | |||||||  |||||||||||||||||
Db        301 ACGGAGTCTCTACAGTCCCTGTGGCCGATTCTTTGCATTTAAGAGTTCTCGCAGCATGG 360

Qy        542 CTGGGCCTAACCTGATCGCAATTGGGTCTTAGTAGTGAATCTCCACAGAGGCCCTTAAGATCA 601
                | || || ||||||| |||||| ||||  || ||||||| |||| |||||||||||||||
Db        361 CCGGACCCAACCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAAGAGCCCTCAAGATCA 420

Qy        602 TGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAA 661
                |||||||||||||||||||| || |||| ||| |||||||| ||| ||||||| || |||
Db        421 TGCAACAGATGAGTGACCATCGTTATGACAAGCTTCACTGTACCCGACGACATGGCCGCCA 480

Qy        662 ACTGTATATATCTAAATATCCCCACAAAAGGGCACGTCTTGCTGCGCACCG-ACCCCGGAAG 721
                |||| |||||||||||||||||| | |||||||  ||||||||||| ||||| |||| |||||
Db        481 ACTGCATATATATCTAAATATCCCCAGCAAAGGGCCATGTCTTGCTGCTGCACCGCACCCCAGAAG 539

Qy        722 AGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGAACCATATGCTGATCCCCGTGA 781
                |||| |||||  |||||||||| ||||||||||| || | || ||| |||||||| ||||
Db        540 AGTACCAGAAGGCCAAGGTCTATGAGAAACTCAAGGCCATCTACTGATCCCTGTGA 599

Qy        782 GCATGTCTGAACTGGAAAGGTGGATGGGCTGCTCACCTGCTGCCTTCAGTTTAAT 836
                ||| | ||| ||||||||||||| ||||||| ||||||||||||||||| ||||||
Db        600 GCAACTCGGAGATGGAAAGGTGGACGGGCTTGCTCACCTGCTGCCTTCACCTGTTTTATT 654
```

RESULT 9

```
302 CGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATCTAAATATC 361
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
682 CCCAACAAAGGCACGTCTTCGTGCCTGACCGACCCCGAAGAGTATCCAGAAAGTGCAAAG 741
    ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
362 CCCAACAAAGGCACGTCTTCGTGCCTGACCGACCCCGAAGAGTATCCAGAAAGTGCAAAG 421
```

Qy

Db

```
742 GTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGAGCATGTGCAAACTGGAAAG 801
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
422 GTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTGCAAACTGGAAAG 481
```

Qy

Db

```
802 GTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAACAA 842
    |||||||||||||||||||||||||||||||||||||||||
482 GTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAACAA 522
```

Qy

Db

RESULT 10

LOCUS   BY754930   698 bp   mRNA   linear   EST 17-DEC-2002
DEFINITION   BY754930 RIKEN full-length enriched, 8 cells embryo Mus musculus
             cDNA clone E860038F15 3', mRNA sequence.
ACCESSION   BY754930
VERSION   BY754930.1 GI:27187153
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 698)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
             Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
             Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
             Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
             Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
             Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
             Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
             Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
             Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
             Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
             Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
             Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
             Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
             Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
             Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
             Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
             Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
             Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
             Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
             Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
             Yang,L., Yuan,J., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
             Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
             Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
             Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
             Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
             Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
             Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
   MEDLINE   22354683
    PUBMED   12466851
  COMMENT   Contact: Yoshihide Hayashizaki
             Laboratory for Genome Exploration Research Group, RIKEN Genomic
             Sciences Center(GSC), Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)
             1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
             Tel: 81-45-503-9222
             Fax: 81-45-503-9216
             Email: genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/
             Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
             Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
             Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
             Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
```

RESULT 9

LOCUS   AI751463   522 bp   mRNA   linear   EST 20-JUN-2002
DEFINITION   AI751463.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
             clone NHTBC_cn10b04 random, mRNA sequence.
ACCESSION   AI751463
VERSION   AI751463.1 GI:5129727
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 522)
  AUTHORS   Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
             Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
             Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
  TITLE   SGAP: The Skeletal Genome Anatomy Project
  JOURNAL   Unpublished (1997)
  COMMENT   Contact: Libin Jia
             Medical Genetics Branch
             National Human Genome Research Institute
             10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
             Tel: 301-402-4877
             Fax: 301-496-7157
             Email: libin@helix.nih.gov
             DNA Sequencing and analyses by National Institutes of Health
             Intramural Sequencing Center (NISC).
             Plate: 10 row: b column: 04
             Seq primer: -21M13 forward primer (ABI).
FEATURES   Location/Qualifiers
     source   1..522
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NHTBC_cn10b04"
                     /sex="Female"
                     /tissue_type="Bone"
                     /cell_type="Trabecular Bone Cells"
                     /lab_host="SURE"
                     /clone_lib="Normal Human Trabecular Bone Cells"
                     /note="Organ: Hip; Vector: pBluescript; Site 1: EcoRI;
                     Library constructed by Dr. Marian Young and Dr. Pamela
                     Gehron Robey (NIDCR)"
ORIGIN

Query Match   60.3%;   Score 517.8; DB 9; Length 522;
Best Local Similarity 99.6%;   Pred. No. 1.7e-86;
Matches 519; Conservative   0; Mismatches   2; Indels   0; Gaps   0;
```
```
322 GCATTAGAAAAACTTCAGTCAATATAGTAGAGATGAGAGATGAAAGATGAAAATGCRACTTTAGAT 381
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
  2 GCACGAGAAAAACTTCAGTCAATATAGTAGAGATGAGAGATGAAAGATGAAAATGCRACTTTAGAT 61

382 GGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTCCAAAAGGACAAAT 441
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 62 GGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAAT 121

442 CAACGAGGTGCTGAAAATCTTGGCTGATACTTTTAAGGACTATGCTATCTCCCAGTGCCA 501
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
122 CAACGAGGTGCTGAAAATCTTGGCTGATACTTTTAAGGACTATGCTATCTCCCAGTGCCA 181

502 GTGGCAGATGGGTTGCATTTGAACGAGTTTTCTGCGACCATGGCCTAACCTGATCCGCA 561
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
182 GTGGCAGATGGGTTGCATTTGAACGAGTTTTCTGCGACCATGGCCTAACCTGATCCGCA 241

562 ATTGGGTCTAGTGAATCTGCACAGAGAGGGCCCTTAAGATCATGCAACAGAGTGACGACCAC 621
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
242 ATTGGGTCTAGTGAATCTGCACAGAGAGGGCCCTTAAGATCATGCAACAGAGTGACGACCAC 301

622 CGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATCTAAATATC 681
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   817 ACCTGCTGCTCAGTTTAATTAACAAGAAGGTAGACTCCTGA 858
         ||||||||||||||||||||||||||||||||||||||||
Db   542 ACCTGCTGCTCCGTTTTATTAACAAGAAGATAGACTCCTGA 583
```

RESULT 11
BB611533

LOCUS   BB611533   731 bp   mRNA   linear   EST 26-OCT-2001
DEFINITION BB611533 RIKEN full-length enriched, 13 days embryo head Mus
            musculus cDNA clone 3110053015 5', mRNA sequence.
ACCESSION   BB611533
VERSION   BB611533.1 GI:16452732
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 731)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE   RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
            Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
            Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES   Location/Qualifiers
     source   1..731
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="3110053015"
                     /tissue_type="head"
                     /dev_stage="13 days embryo"
                     /lab_host="SOLR"
                     /clone_lib="RIKEN full-length enriched, 13 days embryo
                     head"
```

Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

```
FEATURES             Location/Qualifiers
     source          1..698
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="E860038F15"
                     /cell_type="8 cells"
                     /dev_stage="8 cells embryo"
                     /clone_lib="RIKEN full-length enriched, 8 cells embryo"

ORIGIN

Query Match       56.0%;  Score 480.8;  DB 13;  Length 698;
Best Local Similarity   89.0%;  Pred. No. 1.5e-79;
Matches  518;  Conservative   0;  Mismatches   64;  Indels   0;  Gaps   0;

Qy   277 CCCGGGGGCGGCCGAGCCGGAGGAGGTTGACATGATGAAAGAGCATTAGAAAAACTT 336
         ||||||||||| |||||| |||||||||||||||||||||||||| ||||||||||
Db     2 CCCGGGGGCGCCCAGCCGCAGGAAGNAGGTTGACATGATGAAAGAGGCTTTGGAAAAACTT 61

Qy   337 CAGCTCAATATAGTAGGAGATGAAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTTA 396
         ||||||||||||||||||||||||||||||||||||||||| ||||| |||||||||||
Db    62 CAGCTCAACATAGTAGGAGATGAAAGATGAAAATGCAACTTTGGATGGTGTGNGGACGTCCTA 121

Qy   397 TTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGTGCTGAA 456
         ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   122 TTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGAACAAATCAACGAGGTGTCTGAA 181

Qy   457 ATCTTGGCTGATACTTTTAAGGACATATGCAGTCTCCACAGTGCCAGTGGCAGATGGGGTTG 516
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   182 ATCTTGGCTGATACTTTTAAGGACTACACGCAGTCTCCTACAGTCCCTGTGGCCGATTCTTTG 241

Qy   517 CATTTGAAGAGTTTCTGCAGCATGGCCGGACCCACCTGGCCTAACCTGATCGCAATTGGGTTCCAGTGAA 576
         ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||| |||
Db   242 CATTTAAAGAGTTTCTGCAGCATGGCCGGACCCACCTGATTGCAATAGGGTTCCAGCGGA 301

Qy   577 TCTGCACAGAAGGCCCTTAAGATCATCATGCAACAGATGAGTGACCACCGCTACGACAAACTC 636
         ||||||||||||||| |||||||||||||||||||||||||||||||| ||||||||||||
Db   302 TCTGCACAGAAGGCCCTCAAGATCATCAACAGATCATGAGTGACCACCGCTACGACAAGCTC 361

Qy   637 ACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCAACAAAGGGCAC 696
         ||||| |||||||| |||||||| ||||||||||||||| |||||||||| |
Db   362 ACTGTACCCGACGCACATGGCCGCGCCAACTGCATATATCCCCAGCAAAGGGCAT 421

Qy   697 GTCTTGCTGCAGCAGCCCGGAGAGTATCCGAGAAGTGCAAAGGTTTATGAGAAACTG 756
         |||||||||||| |||||||| |||| |||||||||||||||||||||||||||||
Db   422 GTCTTGCTGCAGCAGCCCGAGAGAGTACCCAGAAGCGCAAAGGGTCTATGAGAAACTC 481

Qy   757 AAGGACCCATATGCTGATCCCCGTGACAGTGCATGCTGAACTGGAAAAGGTGGATGGGCTGCTC 816
         |||||||| |||||| |||||||||| || |||||||| |||| |||||||||||||| |||
Db   482 AAGGACCATCTACTGATCCCTGTGAGCAACTCGGAGATGAAAAGGTGGACGGCGGCTTGCTC 541
```

REFERENCE
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
           1 (bases 1 to 800)
           Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
           Geisel,S., Kucaba,T., Lacy,M., Le.M., Martin,J., Morris,M.,
           Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,M.,
           Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
           Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
           WashU-HHMI Mouse EST Project
           Washington University School of MedicineP
           4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
           Tel: 314 286 1800
           Fax: 314 286 1810
           Email: mouseest@watson.wustl.edu
           This clone is available royalty-free through LLNL, contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           MGI:916159
           Possible reversed clone: similarity on wrong strand
           Seq primer: primer name ambiguous
           High quality sequence stop: 488.

FEATURES   Location/Qualifiers
     source   1..800
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1432091"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse kidney mkia"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                     (CACCGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTNGGCCTTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACCGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."

ORIGIN

Query Match   55.0%; Score 472.2; DB 9; Length 800;
Best Local Similarity   86.4%; Pred. No. 6.3e-78;
Matches 579; Conservative   0; Mismatches   84; Indels   7; Gaps   5;

Qy   196 CCGGCCACGAGCAGCCTTCCGACTG--CGTCTTCGTGGGAGGACGTGCGCCGGTGTGCGCG 253
                |||||||||||||||||||||||| |||||||||||||||||||| ||||||||||||
Db   800 CCCGGCCACGAGCAGCCTTGCCCGACTGCGTGTTGTTGAAGAACGTGCCCCGTCGTGTGCG 741

Qy   254 AGGAGA-CGGCCTCCATCACCCGACCCGGG--CGGCCGAGCCGGAGGAAGGA-GGTTGAC 309
                |||| ||||||||||||||||||||||| |||||||||||||||||||| ||||||
Db   740 AGGAGACCGGCCTTCATCACCCGACCCCGGGGCCGCCGAGCCGGAGGAAGGAGGGTTGCC 681

Qy   310 ATGATGAAAGAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAAGATGAAAGAAAAT 369
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   680 ATGATGAAAGAGAGCTTTGAAAAAAACTTCAGCTCAACATAGTAGAAGATGAAAGAAAAT 621

Qy   370 GCAACTTTAGATGGCCGAGATGTTTATTCACAGGCAGAGAATTTT-TGTGGGCCTTTC 428
            ||||||||||||| |||||||||||||||||||||||||||||||| |||||||||||||
Db   620 GCANCTTTGGATGGTGGGACGTCCTATTCACAGGCAGAGAATTTTTNTGTGGGCCTTTC 561

Qy   429 CAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCCAGT 488
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   560 CAAAAGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTTAAGGACTACCCAGT 501

ORIGIN

/note="Site_1: XhoI; Site_2: SstI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAAGGATCCAGAGAGCTCTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 50.0. Second strand
cDNA was prepared with the primer adapter of sequence [5'
GAGAGAGAGATTCTGAGTTAATTAATTAATCCCCCCCCCCC 3']. cDNA
was cleaved with XhoI and SstI."

ORIGIN

Query Match   55.8%; Score 479; DB 10; Length 731;
Best Local Similarity   88.8%; Pred. No. 3.3e-79;
Matches 529; Conservative   0; Mismatches   66; Indels   1; Gaps   1;

Qy   1 ATGGCCGGCTCGGCCACCCTCCGCCTTCGGCCGGCAGCCCACCCACGCCCGTGGTGCGGCG 60
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   134 ATGGCCGGCCCTCGGCCACCCTCCGCCTTCGGCCGGCACCCACGCCCGTGGTGCGGGCG 193

Qy   61 CTACCCGAGTCGCTTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGGAGGAGGTGGACGTC 120
            |||||||||| ||| |||||||||||||| ||||||||||||||||||||||||||
Db   194 CCGCCCGAGTCCTGCGCCACGCGGCTCGCAGGGGAGCGGGCAGGTGGATTTC 253

Qy   121 GCCCGCGCGGAACGGCAGCACCAGCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 180
            ||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db   254 GCTCGCGCCGAGCGCCAGCACGAGCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 313

Qy   181 CAGGTGGTGAGCTGCCGGCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 240
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   314 CAGGTGGTGCAGCTGCCCGACGAGAGCCTTGCGGACTGCGTGTTCGTGGAGGACGTG 373

Qy   241 GCCCGTGGTGCGGAGGAGACGGCCCTCATCACCCGACCCGGGCGGCCGAGCCGGAGGAAG 300
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   374 GCCCGTCGTGCGGAGGAGACGGCCCTTCATCACCCGACCCCGGGGCCGCCCAGCCGGAGGAAG 433

Qy   301 GAGGTTGACATGATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
            |||||||||||||||| |||||| |||||||||||||||||||||||| |||||||||||
Db   434 GAGGTTGACATGATGATGAAAGAGAGCTTTGAAAAAACTTCAGCTCAACATAGTAGAGATGAAA 493

Qy   361 GATGAAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 420
            |||||||||||||||||||||| ||| |||||||||||||||||||||||||||||||
Db   494 GATGAAAAATGCAACTTTGGATGGTGGGACGTCCTATTCACAGGCAGAGAATTTTTTGTG 553

Qy   421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTTAAGGAC 480
            ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   554 GGCCTTTCCAAAAGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTTAAGGAC 613

Qy   481 TATCGAGTCTCCCAGTGCCAGTGCCAGATGGGTTGCATTGAAGAGTTTCTGCAG-CAT 539
            ||||| ||||| |||||||||||||||||| |||||||||||||||| ||||| ||||
Db   614 TACGGAGTCTCTACAGATCCTGTTGGCCGATTCTTGGCCGATTAAAGAGATTCTGCAGCCAT 673

Qy   540 GGCTGGGCCTAACCTGATCGATGCAATTGGGTCTAGTAGAAGTCGACAGAAGGCCCTTA 595
            |||||||||||||||||||| ||||||||||||||||||||||||||||
Db   674 GGCCGGACCCAACCTGATTGCATTAAGGTCCAGCGAATCTGCACAGATGCCCCTTCA 729

RESULT 12
LOCUS   AA986117   800 bp   mRNA   linear   EST 28-MAY-1998
DEFINITION uc81h06.x1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1432091 3' similar to TR:O08557 O08557
            N-G,N-G-DIMETHYLARGININE DIMETHYLAMINOHYDROLASE. ;, mRNA sequence.
ACCESSION   AA986117
VERSION   AA986117.1 GI:3167992
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus

```
Qy  489 CTCCACAGTGCCAGTGCCAGATGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCC 548      Qy  404 GCAGAGAATTTTTGTGGGCCTTTCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGG 463
        ||||| |||||| ||||  ||  |||| |||| |||||||||||||||||||||||||| | |||              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  500 CTCTACAGTGCCCGTGCGCCAATTCTTGCCATTTAAAGAGTTTCTGCAGCATGGCTGGGCGACC 441      Db    1 GCAGAGAATTTTTGTGGGCCTTTCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGG 60

Qy  549 TAACCTGATCGCAATTGGGTCTCAGTGAATCTGCACAGAGGCCCTAAGATCATGCAACA 608      Qy  464 CTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGA 523
        ||||||||||||  |||||||| |  ||  |||||||| |||||| |||||||||||||              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  440 CAACCTGATCGGCAATAGGGTCCAGCAGCAATCTGCACGAAGGCCCTCAAGATCATGCAACA 381      Db   61 CTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGA 120

Qy  609 GATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTAT 668      Qy  524 AGAGTTTCTGCAGCATGGCTGGGCCTAACCGACATCGCAATTGGGTCTCGTAGTGAATCTGCAC 583
        ||||||||||||| || |||||| ||||||| | ||  | |   ||||||||| ||||              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  380 GATGAGTGACCATCGTTATGACAAGCTCACTGACAAGCTGCGCCCCACTGCAT 321      Db  121 AGAGTTTCTGCAGCATGGCTGGGCCTAACCGACATCGCAATTGGGTCTCGTAGTGAATCTGCAC 180

Qy  669 ATATCTAAATATCCCCAACAAAGGGCACGTCTTGCTGCCGAACCCCGGAAGAGTATCC 728      Qy  584 AGAAGGCCCTAAGATCATGCAACAGAGTGAGTGCACACCGCTACGACAAACTCACTGTGC 643
        ||||||||||| || |||||||| |||||| || ||| ||| ||||||| ||||| |||              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  320 ATATCTAAATATCCCCAGCCAAAGGGCATGTCTTGCTGCCGAACCCCGGAAGAGTACCC 261      Db  181 AGAAGGCCCTAAGATCATGCAACAGAGTGAGTGCACACCGCTACGACAAACTCACTGTGC 240

Qy  729 AGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTC 788      Qy  644 CTGATGACATAGCAGCAAACTGTATATATATCTAAATAATCCCAACAAGTTTATGAGAAGTT 703
        ||||||| ||||| || |||||||| ||||| | | |||  ||||| ||||||||  |||  ||              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  260 AGAAAGCGCAAAGGTCATGAGAAACTCAAGGACCATATGCTGATCCCCTGTGAGCAACTC 201      Db  241 CTGATGACATAGCAGCAAACTGTATATATATCTAAATAATCCCAACAAGTTTATGAGAAGTT 300

Qy  789 TGAACTGGAAAAGTGGATGGGTGGCTGCTCACCTGTTTAATTAACAGAAGGT 848      Qy  704 TGCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACC 763
        ||   |  |||||||||||| || ||||| |||||||||  |||||||||||              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  200 GGAGATGGAAAAGGTGGACGGCGCTTGCGGCTTGCTCACCTGCTCCGTTTTATTAACAGAGAT 141      Db  301 TGCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACC 360

Qy  849 AGACTCCTGA 858      Qy  764 ATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGTGGATGGGTGGCTGCTCACCTGCT 823
        ||||||||||              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 AGACTCCTGA 131      Db  361 ATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGTGGATGGGTGGCTGCTCACCTGCT 420

                                                                Qy  824 GCTCAGTTTAATTAACAAGAAGTAGACTCCTGA 858
                                                                        |||||||||||||||||||||||||||||||||
                                                                Db  421 GCTCAGTTTAATTAACAAGAAGTAGACTCCTGA 455
```

RESULT 13
AY406329
LOCUS   AY406329   455 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION Homo sapiens DDAH1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406329
VERSION   AY406329.1 GI:39762303
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 455)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE   2 (bases 1 to 455)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE   Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..455
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene   <1..>455
                     /gene="DDAH1"
                     /locus_tag="HCM2521"
ORIGIN

Query Match   52.8%; Score 453.4; DB 29; Length 455;
Best Local Similarity   99.8%; Pred. No. 1.8e-74;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 14
AY406330
LOCUS   AY406330   455 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION Pan troglodytes DDAH1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406330
VERSION   AY406330.1 GI:39762304
KEYWORDS   GSS.
SOURCE   Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 455)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE   2 (bases 1 to 455)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE   Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..455
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene   <1..>455
                     /gene="DDAH1"
                     /locus_tag="HCM2521"

ORIGIN

```
Query Match          52.5%;    Score 450.2;  DB 29;  Length 455;
Best Local Similarity 99.3%;    Pred. No. 6.9e-74;
Matches    452;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy  404 GCAGAGAATTTTTGTGGGCCTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGG 463
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GCAGAGAATTTTTGTGGGCCTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGG  60

Qy  464 CTGTATACTTTAAGGACTATGCAGTCTCCAGTGCCAGTGCAGATGGGTTGCATTTGA 523
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGTATACTTTAAGGACTATGCAGTCTCCAGTGCCAGTGCAGATGGGTTGCATTTGA 120

Qy  524 AGAGTTTCTGCAGCATGGCTGGGCCTACCTGATCGCAATTCGGTCTAGTGAATCTGCAC 583
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AGAGTTTCTGCAGCATGGCTGGGCCTACCTGATCGCAATTCGGTCTAGTGAATCTGCAC 180

Qy  584 AGAAGGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGC 643
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AGAAGGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTATGACAAACTCACTGTGC 240

Qy  644 CTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAAAGGGACAGTCTTGC 703
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CTGATGACGTAGCAGCAAACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGC 300

Qy  704 TGCACCGAACCCCGGAAGAGTATCCGAAGAAGTGCAAAGTTTATGAGAAACTGAAGGACC 763
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 TGCACCGAACCCCGGAAGAGTATCCGAAGAAGTGCAAAGTTTATGAGAAACTGAAGGACC 360

Qy  764 ATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGTGGAAAAAGTGAACTGCTGCTCACCTGCT 823
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAAGTGGAAAAAGTGAACTGCTGCTCACCTGCT 420

Qy  824 GCTCAGTTTAATTAACAAGAAGTAGACTCCTGA 858
            |||||||||||||||||||||||||||||||||
Db  421 GCTCAGTTTTAATTAACAAGAAGTAGACTCCTGA 455
```

RESULT 15
CB247154

LOCUS   CB247154   685 bp   mRNA   linear   EST 09-JUL-2003
DEFINITION UI-M-FI0-cdy-1-02-0-UI.r1 NIH_BMAP_FI0 Mus musculus cDNA clone
            IMAGE:6836211 5', mRNA sequence.
ACCESSION   CB247154
VERSION   CB247154.1 GI:28368798
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 685)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES   Location/Qualifiers
     source   1..685
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6836211"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 12.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_FI0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
                     Site 2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CAGCCACGAC. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."

ORIGIN

```
Query Match          52.0%;    Score 445.8;  DB 14;  Length 685;
Best Local Similarity 90.1%;    Pred. No. 5.1e-73;
Matches    499;  Conservative   0;  Mismatches  53;  Indels   2;  Gaps   2;

Qy    1 ATGGCCGGCCTTCGGCCACCCCTTCGGCCGGGCCACCCACCGCTGTGCGGGGCG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123 ATGGCCGGCCTTCGGCCACCCCTTCGGCCGGGCCACCCACCGCTGTGCGGGGCT 182

Qy   61 CTACCGAGTCGCTCGCCAGGACGCGCGCTGAGAAGGCCAAGGGCGAGGAGGTGGACGTC 120
            |||||||||||||| |||||||||||||||||||||||| |||||||||||||||
Db  183 CCGCCGAGTCCGTGTGCCGGCGCTGCGAGGGCCTCGCCAGGGGCGAGGAGGTGGATTTC 242

Qy  121 GCCCGGCGGAACGGCCAGCACCAGCTCTACGTGGGCGTGCTGGGCGAGCAAGCTGGGCTG 180
            |||||| |||| ||||||||||||||||||||||||||||||||| |||||||||||||
Db  243 GCTCGCCGGCCGAGCGCCGCCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTG 302

Qy  181 CAGGTGGTGGAGCTGCTGCCGCCGACGAGCCTTCGGACTGCGTCTTCGTCGTGGAGGACGTG 240
            |||||||||||||| ||||| ||||||||||||||| |||||||| |||||||||||||||
Db  303 CAGGTGGTGCAGCTGCTGCCGCCGACGAGCCCGACACCGGGGGCCGGGGCCGGAGGAAG 362

Qy  241 GCCCGTGTGTGCGAGGAGACGCGGCCCTCATCACCGACCACCGCGCCCCCGGGGCGCCGGAGGAAG 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  363 GCCGTCGTGTGCGAGGAGACGGCGCCCTCATCACCGACCACCGCGCCCCCGGGGCGCCAGGAAG 422

Qy  301 GAGGTTGACATGATGAAAGGAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
            |||||||||||||||||||||||||||   |||||||||||||||| || |||||||||||
Db  423 GAGGTTGACATGATGAAAGGAGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAA 482

Qy  361 GATGAAAATGCAACTTT-AGATGGCGGAGATGTTTATTCACAGGCAGAGA-TTTTTG 418
            |||||||||||||||||  |||||||||||| |||||||||||||||||||||| |||||
Db  483 GATGAAAATGCAACTTTGGATGGTGGGGACGTCCTATTCCACAGGCAGAGAATTTTTTG 542

Qy  419 TGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGAATCTTTGGATACTTTTTAAGG 478
            ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  543 TGGGCCTTTCCAAAAGGACAAATCAACGAGGCTGAAATCACGAGGCTGAAATCTTGGCTGATACTTTTNAAGG 602

Qy  479 ACTATGCAGTCTCCACAGTGCCAGTGCAGTGGGTTGCATTGAAGAGTTTCTGCAGCA 538
            |||||||||||| ||||||||||||||||||| |||||||||| ||||||||||||
Db  603 ACTACGCAGTCTCTACAGTCCCTGTGCCGATTCTTTGCATTTNAGAGTTTCTGCAGCA 662

Qy  539 TGGCTGGGCCTAAC 552
            ||||||||||||||
Db  663 TGGCCGGGACCCACC 676
```

Search completed: June 7, 2004, 21:17:39
Job time : 2576 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: June 7, 2004, 16:07:05 ; Search time 97 Seconds
(without alignments)
4908.741 Million cell updates/sec

Title: US-09-889-733B-1
Perfect score: 858
Sequence: 1 atggccggcctcggccaccc.........acaagaggtagactcctga 858

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	185.6	21.6	1397	4	US-09-023-655-37	Sequence 37, Appl
c	2	81.2	9.5	239	4	US-09-702-701-404	Sequence 404, App
c	3	81.2	9.5	239	4	US-09-736-457-404	Sequence 404, App
c	4	81.2	9.5	239	4	US-09-614-124B-404	Sequence 404, App
c	5	81.2	9.5	239	4	US-09-671-325-404	Sequence 404, App
c	6	81.2	9.5	239	4	US-09-589-184-404	Sequence 404, App
	7	69		526	4	US-09-621-976-12670	Sequence 12670, A
	8	55.6	6.5	2598	4	US-09-252-991A-2768	Sequence 2768, Ap
	9	55.6	6.5	2790	4	US-09-252-991A-2864	Sequence 2864, Ap
	10	52.2	6.1	1288	1	US-08-440-856A-9	Sequence 9, Appli
	11	49.8	5.8	1257	4	US-09-252-991A-1225	Sequence 1225, Ap
	12	49.8	5.8	2280	4	US-09-252-991A-1270	Sequence 1270, Ap
	13	49.8	5.8	2616	4	US-09-252-991A-1336	Sequence 1336, Ap
	14	49.6	5.8	7218	1	US-08-232-463-14	Sequence 14, Appl
	15	49.4	5.8	666	6	5472691-4	Patent No. 5472691
c	16	49.4	5.8	669	3	US-08-556-965-1	Sequence 1, Appli
c	17	49.4	5.8	723	6	5472691-7	Patent No. 5472691
c	18	49.4	5.8	1239	4	US-09-252-991A-6622	Sequence 6622, Ap
c	19	49.4	5.8	1389	2	US-08-023-980B-3	Sequence 3, Appli
c	20	49.4	5.8	1389	2	US-08-486-953A-3	Sequence 3, Appli
c	21	49.4	5.8	1396	6	5472691-1	Patent No. 5472691
c	22	49.4	5.8	1449	4	US-09-252-991A-6642	Sequence 6642, Ap
c	23	49.4	5.8	1647	4	US-09-252-991A-6576	Sequence 6576, Ap
c	24	49.4	5.8	1731	4	US-09-252-991A-6675	Sequence 6675, Ap
c	25	49.4	5.8	10079	2	US-08-476-866-20	Sequence 20, Appl
c	26	48.2	5.6	1389	4	US-09-252-991A-4243	Sequence 4243, Ap
c	27	48.2	5.6	2598	4	US-09-252-991A-4323	Sequence 4323, Ap
c	28	48.2	5.6	2874	4	US-09-252-991A-4533	Sequence 4533, Ap
c	29	47.6	5.5	43280	2	US-08-804-227C-1	Sequence 1, Appli
	30	47.2	5.5	47981	4	US-09-679-279-1	Sequence 1, Appli
c	31		47	1074	4	US-09-252-991A-5833	Sequence 5833, Ap
	32		47	1275	4	US-09-252-991A-5813	Sequence 5813, Ap
	33		47	1554	1	US-08-440-856A-5777	Sequence 5777, Ap
	34	46.6	5.4	1187	1	US-08-440-856A-2	Sequence 2, Appli
	35	46.4	5.4	519	4	US-09-252-991A-16355	Sequence 16355, A
	36	46.4	5.4	609	4	US-09-252-991A-16464	Sequence 16464, A
	37	46.4	5.4	4257	2	US-08-690-473-1	Sequence 1, Appli
	38	46.4	5.4	4257	3	US-09-259-821A-1	Sequence 1, Appli
	39	46.4	5.4	4257	1	US-08-843-659-1	Sequence 1, Appli
c	40	46.4	5.4	12001	1	US-08-458-568A-11	Sequence 11, Appl
c	41	46.2	5.4	1926	4	US-09-249-585A-4	Sequence 4, Appli
c	42	46.2	5.4	1931	2	US-09-130-114-2	Sequence 2, Appli
	43	46.2	5.4	11220	3	US-09-105-537-32	Sequence 32, Appl
	44	46.2	5.4	15664	1	US-08-402-282-3	Sequence 3, Appli
	45	46.2	5.4	15664	1	US-08-508-004-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-023-655-37

; Sequence 37, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09,023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1397 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 003437
US-09-023-655-37

Query Match   21.6%; Score 185.6; DB 4; Length 1397;

```
Best Local Similarity   57.4%;   Pred. No. 4.7e-35;   Indels   2;   Gaps   2;
Matches   372;   Conservative   0;   Mismatches   274;

Qy   29  TCGCCGGCGCCACCACGCCGTGTGCGGCGCCTACCGAGTCGCTCTGCCAGCACGCGC        88
             |||   |||||  |||||||  ||    ||||||||||||  ||| ||||||||||
Db  357  TGGGCCGCTGCTCCCATGCCGGGGGGAGTCCCAGAGAGCCTGGCGTCGGGGGAAG         416

Qy   89  TGAGAAGGCCAAGGGCGAGGAGGTCGCTCGCCGCGCGGAACGGCAGCACCAGCTCT        148
Db  417  GTGCGGGGGCTGGCCTTCCGCTCTGGATCTGCCAAAGCTCAAAGGGAGCACGGGGTGC      476

Qy  149  ACGTGGGCGGTGCTGGCGAGCAAGCTGGGGCTGCAGGTGGTGGCGCCGCGACGAGA        208
Db  477  TGGGAGGTAAACTGAGGCAACGACTGGGCTACAGCTGCTACGAACTGCCACCTGAGGAGT    536

Qy  209  GCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACGGCCCTCA    268
Db  537  CATTGCCGCGCTGCCTGGGCACGGCCGTGATCCAAGGGGACACGGCCCTAA             596

Qy  269  TCACCCGACCCGGGGGCCGCCGGAGGCCGGAGGAGGAGGTTGACATGATGAAAGAAGCATTAG  328
Db  597  TCACGGCCGGCCCTGGAGCCCGCTCGTAGGCCAGAGGTCGATGGAGTCCGAAGCCCTGC      656

Qy  329  AAAAACTTCAGCTCAATATAGTAGGATGAAAGATGAAAATGCAACTTTAGATGGGCCGGAG    388
Db  657  AAGACCTGGGGCTCCGAATTGTGGAAATAGGAGACGACGCCGACGCTGGATGGCACTG       716

Qy  389  ATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCAAAGGACAAATCAACGAG        448
Db  717  ACGTTCTCTTCACCGGCGCGGGAGTTTTTCGTAGGCCTCTCCAAATGACCAATCACCGAG     776

Qy  449  GTGCT-GAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCA    507
Db  777  GCGCTGGAGATCGTGGGGGACACGTTCCGGGGACTTCGCCGTCTCCACTGTGCCAGTCTCG    836

Qy  508  GATGGGTTGCCATTTGAAGAGTTTCTCGCAGCATGGCTGGGGCCTAACCTGATCGCAATTGGG   567
Db  837  GGTCCCTCCACCTGCGCGGGTCTCGTGCGCGATGCGGGGGACCCCGCACTGTTGTGCGGAGGC    896

Qy  568  TCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCCTAC      627
Db  897  AGCAGCGACGCTGCCAAAAGGCTGTCCCGGGCAATGGCAGTGCTGACAGATCACAGATCACCCATAT 956

Qy  628  GACAAACTCA-CTGTGCCTGATGACATAGCAGCAAACTGTATATATCT                674
Db  957  GCCTCCCTGACCCCTCCCAGATGACGGACGTGCTGCTGACTGTCTTTCT               1004
```

RESULT 2
US-09-702-705-404/c
; Sequence 404, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien

US-09-702-705-404

```
Query Match         9.5%;   Score 81.2;   DB 4;   Length 239;
Best Local Similarity   58.8%;   Pred. No. 1.9e-10;
Matches   140;   Conservative   0;   Mismatches   98;   Indels   0;   Gaps   0;

Qy  185  TGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCG     244
             |||| |   ||||||   ||  ||| ||||||     ||||| ||
Db  238  TGCTAGAAACTGCCACCTGAGGAGTCATTGCCGGCCGTGCTGGGACCGCTGCTTGGCGGACCACGGCCG  179

Qy  245  TGGTGTGCGAGGAGACGGCCCTCATCACCCGGGGGCCCGAGGAGGAGGAGGAGG           304
Db  178  TGATCCAAGGGGACACGGCCCTAATCACGGCGCCCTGGAGCCCGCTCGTAGGCCAGAGG       119

Qy  305  TTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGGATGAAAGATG      364
Db  118  TCGATGGAGTCCGCAAAGCCCTGCAAGACCTGGGGCTCCGAATTGTGGAAATAGGAGACG      59

Qy  365  AAAATGCAACTTTAGATGGGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGG          422
Db   58  AGAACCGCACCTGGATGGCACTGACGTTCTCTTCACCGGCGCGGGGAGTTTTTCGTAGG       1
```

RESULT 3
US-09-736-457-404/c
; Sequence 404, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien

US-09-736-457-404

```
Query Match         9.5%;   Score 81.2;   DB 4;   Length 239;
Best Local Similarity   58.8%;   Pred. No. 1.9e-10;
Matches   140;   Conservative   0;   Mismatches   98;   Indels   0;   Gaps   0;

Qy  185  TGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCG     244
Db  238  TGCTAGAAACTGCCACCTGAGGAGTCATTGCGGSTGGGACCGCTGCTTGGCGGACCACGGCCG   179

Qy  245  TGGTGTGCGAGGAGACGGCCCTCATCACCCGGGGGCCCGAGGAGGAGGAGGAGG           304
Db  178  TGATCCAAGGGGACACGGCCCTAATCACGGCGCCCTGGAGCCCGCTCGTAGGCCAGAGG       119

Qy  305  TTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGGATGAAAGATG      364
Db  118  TCGATGGAGTCCGCAAAGCCCTGCAAGACCTGGGGCTCCGAATTGTGGAAATAGGAGACG      59

Qy  365  AAAATGCAACTTTAGATGGGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGG          422
Db   58  AGAACCGCACCTGGATGGCACTGACGTTCTCTTCACCGGCGCGGGGAGTTTTTCGTAGG       1
```

RESULT 4

US-09-614-124B-404/c
; Sequence 404, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C9
;   CURRENT APPLICATION NUMBER: US/09/614,124B
;   CURRENT FILING DATE: 2001-07-11
;   NUMBER OF SEQ ID NOS: 1668
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
;   LENGTH: 239
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-614-124B-404

Query Match   9.5%;   Score 81.2;   DB 4;   Length 239;
Best Local Similarity 58.8%; Pred. No. 1.9e-10;
Matches 140; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy   185 TGGTGGAGCTGCCGGCCAGCAGAGAGCCTTCCGACTGCGTCTTCGTGGAGGACGTGGCCG 244
Db   238 TGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCGCTGCTTGGCGACACGGCCG 179

Qy   245 TGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCGAGGAAGGAGG 304
Db   178 TGATCCAAGGGGACACGGCCCTAATCACGCGGCCCTGGAGCCCCGCTCGTAGGCCAGAGG 119

Qy   305 TTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATG 364
Db   118 TCGATGGAGTCCCGCAAAGCCCTGCAAGACCTGGGGCTCCGAATTGTGGAAAATAGGAGACG 59

Qy   365 AAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTGGG 422
Db   58   AGAACGCGACGCGGATGGCACTGACGTTCTCTTCACCGGCCGGGAGTTTTTCGTAGG 1

RESULT 5
US-09-671-325-404/c
; Sequence 404, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   APPLICANT: Fan, Liqun
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C12
;   CURRENT APPLICATION NUMBER: US/09/671,325
;   CURRENT FILING DATE: 2000-09-26
;   NUMBER OF SEQ ID NOS: 1825
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
;   LENGTH: 239
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-671-325-404

Query Match   9.5%;   Score 81.2;   DB 4;   Length 239;
Best Local Similarity 58.8%; Pred. No. 1.9e-10;
Matches 140; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy   185 TGGTGGAGCTGCCGGCCAGCAGAGAGCCTTCCGACTGCGTCTTCGTGGAGGACGTGGCCG 244
Db   238 TGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCGCTGCTTGGCGACACGGCCG 179

Qy   245 TGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCGAGGAAGGAGG 304
Db   178 TGATCCAAGGGGACACGGCCCTAATCACGCGGCCCTGGAGCCCCGCTCGTAGGCCAGAGG 119

Qy   305 TTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATG 364
Db   118 TCGATGGAGTCCCGCAAAGCCCTGCAAGACCTGGGGCTCCGAATTGTGGAAAATAGGAGACG 59

Qy   365 AAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTGGG 422
Db   58   AGAACGCGACGCGGATGGCACTGACGTTCTCTTCACCGGCCGGGAGTTTTTCGTAGG 1

RESULT 6
US-09-589-184-404/c
; Sequence 404, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C8
;   CURRENT APPLICATION NUMBER: US/09/589,184
;   CURRENT FILING DATE: 2000-06-05
;   NUMBER OF SEQ ID NOS: 827
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
;   LENGTH: 239
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-589-184-404

Query Match   9.5%;   Score 81.2;   DB 4;   Length 239;
Best Local Similarity 58.8%; Pred. No. 1.9e-10;
Matches 140; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy   185 TGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGACTGCGTCTTCGTGGAGGACGTGGCCG 244
Db   238 TGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCGCTGCTTGGCGACACGGCCG 179

Qy   245 TGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCGAGGAAGGAGG 304
Db   178 TGATCCAAGGGGACACGGCCCTAATCACGCGGCCCTGGAGCCCCGCTCGTAGGCCAGAGG 119

Qy   305 TTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATG 364
Db   118 TCCATGGAGTCCCGCAAAGCCCTGCAAGACCTGGGGCTCCGAATTGTGGAAATAGGAGACG 59

Qy   365 AAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTGGG 422
Db   58   AGAACGCGACGCTGGATGGCACTGACGTTCTCTTCACCGGCCGGGAGTTTTTCGTAGG 1

RESULT 7
US-09-621-976-12670
; Sequence 12670, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:

```
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 12670
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-12670

Query Match             8.0%;  Score 69;  DB 4;  Length 526;
Best Local Similarity   52.2%;  Pred. No. 1.9e-07;
Matches   153;  Conservative  0;  Mismatches  140;  Indels  0;  Gaps  0;

Qy   88  CTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTCGCCCCGCGGAACGGCGCAGCACCAGCTC  147
             |||||  |||    ||||    || |     ||||||||  | ||  | || ||||
Db  234  CTAAGAAATCCGAGCGGTTGGAGGGGAGTCGTGTGGGATGGGATGGGACGCCGGGGTG    293

Qy  148  TACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGGTGGAGCTGCCGGCCGACGAG       207
             |||  ||||       || ||||| |||   |||     ||  ||||  ||| |
Db  294  CTGGGAGGTAAACTGAGCGAACGACTGGGGCTACAGCCTGGGGCTAGAACTGCCACCTGAGGAG  353

Qy  208  AGCCTTCCGGACTGCCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACGGCCCTC   267
             || || ||||  ||| ||     ||||  | |||||||||
Db  354  TCATTGCCGCTGGGGACCGCGTCCTTGGCGCAACGCGCCGTGATCCAAGGGGACACGGCCCCTA  413

Qy  268  ATCACCCGACCCCGGGGCGCCGAGCCGGAGGAGGTTGACATGATGAAAGAGCATTA        327
             ||||    ||||| ||  ||||    |||     || ||||| | ||||| | |||
Db  414  ATCACGGGCCCTGGAGCCCCGTCGTAGGCCAGAGGTCCCCAAAGCCCCTG             473

Qy  328  GAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGA           380
             |||   ||||||||     |||    ||||    ||||||||  ||||||| |
Db  474  CAAGACCTGGGCTCGGAATTGTGAAAATAGGAGACGAGAACGCACGCTGGA            526
```

```
RESULT 8
US-09-252-991A-2768
; Sequence 2768, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2768
; LENGTH: 2598
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2768

Query Match             6.5%;  Score 55.6;  DB 4;  Length 2598;
Best Local Similarity   49.2%;  Pred. No. 0.00052;
Matches   176;  Conservative  0;  Mismatches  179;  Indels  3;  Gaps  1;

Qy   83  ACGCGCTGAGAAGCGCGAGGAGGTGGAGGTGGACGTCGCCCCGCGGCGCGGAACGGCGCAGCACC   142
Db 2072  ACGGCCTGACCTCCAGCCACCTCGGGCAGCCACCTGGGGCTGGGCAGCAAGCCCTGGACCAGCAGCACA 2131

Qy  143  AGTCTTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGGTGGACATGGACACCCGTGCCGCCG    202
Db 2132  ACGCCTACATCCGGCCCTTGCAGACC---TGCGACGTGGACATCACCCGTGCGCCGG         2188

Qy  203  ACGAGAGCCTTCCGGACTGCGTCTTCGTTGGAGACGTGGCCGTGGTGTGCGAGGAGACGG    262
Db 2189  ACGAAGCGTTCCCCGACTCGGTGTTCGTCGAGGAACGGAGGACGGCCGGGCGGCCGAGGAGAAA 2248

Qy  263  CCCTCATCACCCGACCCCGGGGCGCCGAGCCGGAGGAGGTTGACATGATGAAAGAGACGG   262
Db 2249  CCATCATCACCCGCCCCCGCGCCGAATCGGCGGCGCGGAGATCATCGGGAAA          2308

Qy  323  CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAATGAACTTTAGATG          382
Db 2309  CCGTGCAGCGCTTCTATCCGGCAAGGTCGGGACCATCGAGCGA                  2368

Qy  383  GCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGGCCTTTCCAAAAGGACAAA      440
Db 2369  CCGGCGGACATCATGATGGTCGGGCGACCACTTCTACATCGGCCCCGCACCAA         2426
```

```
RESULT 9
US-09-252-991A-2864
; Sequence 2864, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2864
; LENGTH: 2790
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2864

Query Match             6.5%;  Score 55.6;  DB 4;  Length 2790;
Best Local Similarity   49.2%;  Pred. No. 0.00053;
Matches   176;  Conservative  0;  Mismatches  179;  Indels  3;  Gaps  1;
```

```
RESULT 10
US-08-440-856A-9
; Sequence 9, Application US/0844085 6A
; Patent No. 5750873
```

GENERAL INFORMATION:
;   APPLICANT: DELLAPORTA, STEPHEN L.
;   TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING
;   TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS
;   NUMBER OF SEQUENCES: 9
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORRISON & FOERSTER
;   STREET: 2000 PENNSYLVANIA AVE. N.W.
;   CITY: WASHINGTON
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20037
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/440,856A
;   FILING DATE: 15-MAY-1995
;   CLASSIFICATION: 800
;   ATTORNEY/AGENT INFORMATION:
;   NAME: MILLMAN, ROBERT A.
;   REGISTRATION NUMBER: 36,217
;   REFERENCE/DOCKET NUMBER: 05463-20001.00
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 887-1517
;   TELEFAX: (202) 887-0763
;   TELEX: 706141
;   INFORMATION FOR SEQ ID NO: 9:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1288 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: both
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-440-856A-9

Query Match   6.1%; Score 52.2; DB 1; Length 1288;
Best Local Similarity   41.8%; Pred. No. 0.0026;
Matches 105; Conservative 29; Mismatches 117; Indels 0; Gaps 0;

Qy   11 TCGGCCACCCTCCGCCTTCGGCCGCGGCCACCACGCCGTGGTGCGGCGCTACCCGAGT 70
            ||| :|||| ||| || | || ||| : | ||| || | | ||||
Db   343 TCGCGGACATCGACGACGCSGCGGGSGAGCGCGTTGGGCKCGGCGCCTGGGSCCCCASGTCR 402

Qy   71 CGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTCGCCCGCGCGG 130
            | : ||||||| | | : | | | ::::||||| | | :::|| | | |||
Db   403 GSTTCGTCGGSTCGACGTGTCSGTGAGGASGACGTSSRGCGCGCCGTSGACTGSGCGC 462

Qy   131 AACGGCAGCACCAGCTTCACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGGTGGTGG 190
            | |::: : | ::: || | |:| || |||| ||| |||
Db   463 YGTSGCGCSRYGRCGGSCGSCCTSGACGTSYWSTGCAACAACGCCGGGTGCTGGGCGCC 522

Qy   191 AGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGT 250
            || ||||||||| || || ||| | | || || |||| ||| || || ||
Db   523 AGACGGCGCGGGCCARGAGCATCCTGTCSTTCGACGCGSGSGAGTTCGACGCGGTGCTCC 582

Qy   251 GCGAGGAGACG 261
            ||| ||| |
Db   583 GCGGTCAACGCG 593

RESULT 11
US-09-252-991A-1225
; Sequence 1225, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;   APPLICANT: Marc J. Rubenfield et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;   TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 107196.136
;   CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 1225
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1225

Query Match   5.8%; Score 49.8; DB 4; Length 1257;
Best Local Similarity   48.7%; Pred. No. 0.0097;
Matches 135; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy   2 TGGCCGGCCTCGGCCACCCTCCGCCTTCGGCCGCCGGCCACCCACGCCGTGGTGCGGGCGC 61
            ||||| || || ||||| | |||||||||| | ||| |||||||||| || ||
Db   717 TGGTCGACATCGGCAACCGCTGCGCGACCGGCAAGGACCTGAGCACCGTGCTCGACGTGA 776

Qy   62 TACCCGAGTCGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTCG 121
            || ||||| ||||||| | | ||| | ||| | || ||||| || | || | |
Db   777 TGAACCAGTCGATCACCCTGATGAACCCGGAGAACTCCGAGCCCAGCCAGCTGGCCACGCTCA 836

Qy   122 CCCGCGCGGAACGGCAGCGAGCACCAGCTTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGC 181
            ||||| || | ||| ||| | ||| | | | ||| ||| | || || |||||| ||||
Db   837 AGCCCGCGACGTGCTGATCCAGCCGCGCGCCGCGTGTCCGGCTACGGCCACCACCGACTTCCGCC 896

Qy   182 AGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGG 241
            || | |||| | ||| ||| ||| |||| ||| || || | || ||| |||| ||
Db   897 GCGTGCCGCAACTGATCGACGGCGGCTACCGCGCCGCCACCGTGCTCGCCGCGCGCCTCG 956

Qy   242 CCGTGGTGCGGAGGAGACGGCGCCCCTCATCACCCGACC 278
            |||| || || | || ||| || | ||| ||| ||
Db   957 CCGAGTTGCGGAAGCCGAAGGACCTCAACAGCGAGC 993

RESULT 12
US-09-252-991A-1270
; Sequence 1270, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;   APPLICANT: Marc J. Rubenfield et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;   TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 107196.136
;   CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 1270
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1270

Query Match   5.8%; Score 49.8; DB 4; Length 2280;
Best Local Similarity   48.7%; Pred. No. 0.012;
Matches 135; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy   2 TGGCCGGCCTCGGCCACCCTCCGCCTTCGGCCGCCGGCCACCCACGCCGTGGTGCGGGCGC 61
            ||||| || || ||||| | |||||||||| | ||| |||||||||| || ||
Db   779 TGGTCGACATCGGCAACCGCTGCGCGACCGGCAAGGACCTGAGCACCGTGCTCGACGTGA 838

Qy   62 TACCCGAGTCGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTCG 121
            || ||||| ||||||| | | ||| | ||| | || ||||| || | || | |
Db   839 TGAACCAGTCGATCACCCTGATGAACCCGGAGAACTCCGAGCCCAGCCAGCTGGCCACGCTCA 898

Qy   122 CCCGCGCGGAACGGCAGCAGCGCCACCAGCTTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGC 181
            ||||| || || ||| || | ||| | | | || ||| ||| | || || |||||| ||||

RESULT 13
US-09-252-991A-1336/c
; Sequence 1336, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1336
; LENGTH: 2616
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1336

Query Match   5.8%;   Score 49.8; DB 4;   Length 2616;
Best Local Similarity 48.7%;   Pred. No. 0.012;
Matches 135; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy   2 TGGCGGGCGCCTCGGCCACCCTCCCGCCTTCGGCCGGGCCACCCACGCCGTGGTGCGGGCGC 61
Db   1517 TGGTCGACATCGGCAACATGGCCGCTGGCGGCCGAGGGGACCTGACACCCGTGCTCGACGTGA 1458

Qy   62 TACCCGAGTCGCTCTGCCAGCACGCGGCTGAGAGCGCCAAGGGCCGAGAGGGTGGACGTCG 121
Db   1457 TGAACCAGTCGATCACCCTGATGACCCGGAAGAACTCCGAAGCCCAGCTGGCCACGCTCA 1398

Qy   122 CCCGCGGGAACGGCAGCGACCACACCAGCTCTACGTGGCGTGTGCTGGGCAGCAAGCTGGGCTGC 181
Db   1397 AGCCCGGCGACGTGCTGATCCAGCCGCCGCGTGTCCGGCTACGGCACCACCGACTTCGGCC 1338

Qy   182 AGTGGTGGAGCTGCCCGGCCGACGAGAGCCTTCCGGACTGCCGTCTTCGTGGAGGACGTGG 241
Db   1337 GGCGTGCCCGCAACTGACGATCGACGCGCGGCTACCGCCGCCACCACGGGTGCTCGCCCGCGCCCTCG 1278

Qy   242 CCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACC 278
Db   1277 CCGAGTTCGCGCAAGCCGAAGGACCTCAACAGCGAAGC 1241

RESULT 14
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pT2gpt-F1s
US-08-232-463-14

Query Match   5.8%;   Score 49.6; DB 1;   Length 7218;
Best Local Similarity   1.1%;   Pred. No. 0.019;
Matches   4; Conservative 222; Mismatches 146; Indels 0; Gaps 0;

Qy   88 CTGAGAAGCGCAAGGCGAGGAGGTGGACGTCGCCCGCGGAACGGCAGCGGCAGCCACCAGCTC 147
Db   1435 CRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1376

Qy   148 TACGTGGCGTGCTGGGCAGCAAGTCGGGCTCGCAGGTGGTGGAGCTGCCGGCCGACGAG 207
Db   1375 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1316

Qy   208 AGCCTTCCGGACTGCCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACGGCCCTC 267
Db   1315 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1256

Qy   268 ATCACCCGACCCGGGGCGCCGAGGAAGGAGGTTGACATGATGAAAGAAGCATTA 327
Db   1255 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1196

Qy   328 GAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATGCGGA 387
Db   1195 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1136

Qy   388 GATGTTTTATTCACAGGCAGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAAATCAACGA 447
Db   1135 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1076

Qy   448 GGTGCTGAAATC 459
Db   1075 RRRRRRRRRATC 1064

RESULT 15
5472691-4/c
; Patent No. 5472691
; APPLICANT: MARKLUND, STEFAN;EDLUND, THOMAS
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE
; NUMBER OF SEQUENCES: 7

Query Match   5.8%;   Score 49.4; DB 6; Length 666;
Best Local Similarity   50.2%;   Pred. No. 0.0098;
Matches 122; Conservative   0; Mismatches 121; Indels   0; Gaps   0;

Qy   23 CGGCCTTCGGCCGGGCCACCCACGCCGTGGTGCGGCGCTACCCGAGTCGCTCTGCCAGC 82
               | ||||||||| |||| | | | | |
Db   593 CCGGGCCCGCACACGCCCACCACGCAGCAGGCCAGCCGCCCGGCCCGCGTTCCCGTTCTCC 534

Qy   83 ACGCGCTGAGAAGCGCCAAGGGCGCGAGGAGGTGGACGTCGCCCGCGGGAACGGCAGCACC 142
               | ||| | | |||| | |||| | | |||||| | | ||
Db   533 ACGGCTGGCCTGGTTGCCCGCGCGGGCCCAGGTCGTCCTCGCCAGCGTGGACGGACCACGGCC 474

Qy   143 AGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCCGCCG 202
               | ||| | | | || | | | |||| | |
Db   473 CGGCCCACGATGGAGTGCGGGCCCGCAGGCGGCCAGGCCGCCAGCCGCGCGTACCTCCAG 414

Qy   203 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACAGTGGCCGTGGTGCGAGGAGACGG 262
               ||||| | | ||| | | | |||| | | | | |
Db   413 AGGCTGCCGTCGCGTCGCGGACCGCGAAGTTGCCGAAGTGCCCCGGGTGCTGCGGGTGCGGCACG 354

Qy   263 CCC 265
               |
Db   353 GCC 351

Search completed: June 7, 2004, 20:34:32
Job time : 100 secs

```
OM nucleic - nucleic search, using sw model

        Copyright (c) 1993 - 2004 Compugen Ltd.
        GenCore version 5.1.6

Run on:    June 7, 2004, 19:33:33 ; Search time 417 Seconds
                                     (without alignments)
                          9386.553 Million cell updates/sec

Title:          US-09-889-733B-1
Perfect score:  858
Sequence:       1 atggccggctcggccaccc.........acaagaggtagactcctga 858

Scoring table:  IDENTITY.NUC
                Gapop 10.0 , Gapext 1.0

Searched:       295936 seqs, 228098010 residues

Total number of hits satisfying chosen parameters:  5991872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :     Published Applications NA:*
         1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
         2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
         3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
         4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
         5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
         6: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
         7: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
         8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
         9: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
        10: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
        11: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
        12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
        13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
        14: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
        15: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
        16: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
        17: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
        18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
        19: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
        19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.		Score	% Match	Query Match	Length	DB	ID	Description
1		851.6	99.3	1633	17	US-10-450-826-9	Sequence 9, Appli
2		851.6	99.3	4208	17	US-10-468-125-22	Sequence 22, Appl
3		715.6	83.4	3008	15	US-10-205-219-83	Sequence 83, Appl
4		705.4	82.2	3778	15	US-10-252-157-328	Sequence 328, App
5		630.4	73.5	3651	9	US-09-252-300-512	Sequence 512, App
6		197.8	23.1	770	9	US-09-910-943-9	Sequence 9, Appli
7		185.6	21.6	1397	17	US-10-641-643-37	Sequence 37, Appl
8		156	18.2	332	9	US-09-783-590-5309	Sequence 5309, Ap
9		148.4	17.3	472	10	US-09-918-995-9117	Sequence 9117, Ap
c 10		148	17.0	83946	17	US-10-450-826-8	Sequence 8, Appli
11		146	17.0	452	10	US-09-918-995-14871	Sequence 14871, A
12		134.2	15.6	403	10	US-09-918-995-19254	Sequence 19254, A
c 13		119.4	13.9	3189	10	US-09-814-353-20553	Sequence 20553, A
c 14		81.2	9.5	239	9	US-09-736-457-404	Sequence 404, App
c 15		81.2	9.5	239	9	US-09-902-941-404	Sequence 404, App
c 16		81.2	9.5	239	9	US-09-849-626-404	Sequence 404, App
c 17		81.2	9.5	239	10	US-09-476-300-404	Sequence 404, App
c 18		81.2	9.5	239	13	US-10-283-017-404	Sequence 404, App
c 19		81.2	9.5	239	15	US-10-017-754-404	Sequence 404, App
c 20		81.2	9.5	239	15	US-10-113-872-404	Sequence 404, App
c 21		65.2	7.6	774	15	US-10-156-761-1679	Sequence 1679, Ap
22		65.2	7.6	9025608	15	US-10-156-761-1	Sequence 1, Appli
23		60.2	7.0	975	13	US-10-107-431-14	Sequence 14, Appl
24		60.2	7.0	45055	13	US-10-107-431-277	Sequence 277, App
25		60	7.0	60	10	US-09-908-975-9030	Sequence 9030, A
26		56.8	6.6	1647	13	US-10-425-114-28074	Sequence 28074, A
27		55	6.4	3729	13	US-10-671-403-86	Sequence 86, Appl
28		55	6.4	3729	13	US-10-671-419-86	Sequence 86, Appl
29		55	6.4	3729	13	US-10-670-844-86	Sequence 86, Appl
30		55	6.4	3729	13	US-10-671-134-86	Sequence 86, Appl
31		55	6.4	3729	13	US-10-673-098-86	Sequence 86, Appl
32		55	6.4	3729	17	US-10-672-638-86	Sequence 86, Appl
33		55	6.4	3729	17	US-10-673-127-86	Sequence 2176
34		52.4	6.1	3012	17	US-09-738-626-2176	Sequence 2176
c 35		52.4	6.1	3309400	9	US-09-738-626-1	Sequence 1, Appli
36		52	6.1	903	9	US-09-815-242-4130	Sequence 4130, Ap
37		52	6.1	903	13	US-10-282-122A-7433	Sequence 7433, Ap
38		52	6.1	2385	16	US-10-369-493-31665	Sequence 31665, A
39		51	5.9	2073	12	US-09-758-759-158	Sequence 158, App
c 40		51	5.9	7788	15	US-10-329-079-8	Sequence 8, Appli
c 41		51	5.9	37360	15	US-10-329-079-6	Sequence 6, Appli
42		51	5.9	109519	12	US-09-758-759-1	Sequence 1, Appli
43		50.8	5.9	585	15	US-10-156-761-6221	Sequence 6221, Ap
44		50.8	5.9	1749	13	US-10-282-122A-25460	Sequence 25460, A
c 45		50.8	5.9	9025608	15	US-10-156-761-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-10-450-826-9
; Sequence 9, Application US/10450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
; APPLICANT:  Ji, Darren
; APPLICANT:  Axelrod, Douglas W.
; APPLICANT:  Cook, Jonathon S.
; APPLICANT:  Jaiswal, Neelam
; APPLICANT:  Eistein, Richard
; APPLICANT:  Houghton, Adam
; APPLICANT:  Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiatic
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 9
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AB001915
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1633)
; OTHER INFORMATION: n = a or c or g or t
US-10-450-826-9

  Query Match             99.3%;  Score 851.6;  DB 17;  Length 1633;
  Best Local Similarity   99.5%;  Pred. No. 6.5e-235;
  Matches  854;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LU, Yan
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SANJANWALA, Madhusudan
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh

Query Match   99.3%;   Score 851.6;   DB 17;   Length 4208;
Best Local Similarity 99.5%;   Pred. No. 1.1e-234;
Matches 854;   Conservative 0;   Mismatches 4;   Indels 0;   Gaps 0;

Qy   1 ATGGCCGGCCTCGGCCACCCTCCGGCTTCGGCCGGCACCCACGCCGTGGTGCGGGCG 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 368 ATGGCCGGGCTCGGCCACCCGCCGGCTTCGGCCGGCACCCACGCCGTGGTGCGGGCG 427

Qy   61 CTACCCGAGTCGCTCGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 428 CTACCCGAGTCGCTCGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 487

Qy 121 GCCCGCGCGGAACGGCAGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 488 GCCCGCGCGGAACGGCAGCAGCACCAGCTCTACGTGGGGCGTGCTGGGCAGCAAGCTGGGGCTG 547

Qy 181 CAGGTGGTGGAGCTGCCGGCCGACGAGCAGCCTTCCGGACTGCGTCTTCGTGGGAGGACGTG 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 548 CAGGTGGTGGAGCTGCCGGCCGACGAGCAGCCTTCCGGACTGCGTCTTCGTGGGAGGACGTG 607

Qy 241 GCCGTGGTGTGCGAGGAGACGCGCCCTCATCACCCGACCCCGAGCCCGAGCCGGAGGAAG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 608 GCCGTGGTGTGCGAGGAGACGGCGCCCTCATCACCCGACCCCGAGCCCGAGCCGGAGGAAG 667

Qy 301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 668 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 727

Qy 361 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420

; APPLICANT: ASTROMOFF, Anna
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: DING, Li
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: FORSYTHE, Ian J.

Qy   1 ATGGCCGGCCTCGGCCACCCTCCGGCTTCGGCCGGCACCCACGCCGTGGTGCGGGCG 60
Db 323 ATGGCCGGGCTCGGCCACCCGCCGGCTTCGGCCGGCACCCACGCCGTGGTGCGGGCG 382

Qy   61 CTACCCGAGTCGCTCGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 120
Db 383 CTACCCGAGTCGCTCGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 442

Qy 121 GCCCGCGCGGAACGGCAGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 180
Db 443 GCCCGCGCGGAACGGCAGCAGCACCAGCTCTACGTGGGGCGTGCTGGGCAGCAAGCTGGGGCTG 502

Qy 181 CAGGTGGTGGAGCTGCCGGCCGACGAGCAGCCTTCCGGACTGCGTCTTCGTGGGAGGACGTG 240
Db 503 CAGGTGGTGGAGCTGCCGGCCGACGAGCAGCCTTCCGGACTGCGTCTTCGTGGGAGGACGTG 562

Qy 241 GCCGTGGTGTGCGAGGAGACGCGCCCTCATCACCCGACCCCGAGCCCGAGCCGGAGGAAG 300
Db 563 GCCGTGGTGTGCGAGGAGACGGCGCCCTCATCACCCGACCCCGAGCCCGAGCCGGAGGAAG 622

Qy 301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
Db 623 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 682

Qy 361 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420
Db 683 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 742

Qy 421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
Db 743 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 802

Qy 481 TATGCCAGTCTCCACAGTGCCAGTGCCAGTGGGTGCATTTGAAGAGTTTCTGCAGCATG 540
Db 803 TATGCAGTCTCCACAGTGCCAGTGCCAGTGGGTTGCATTTGAAGAGTTTCTGCAGCATG 862

Qy 541 GCTGGGCCTAACCTGCAATTGGGTCTTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600
Db 863 GCTGGGCCTAACCTGCAATTGGGTCTTAGTGAATCTGCACAGAAGGCCCTTAAGATC 922

Qy 601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGCCTGATGACATAGCAGCA 660
Db 923 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGCCTGATGACATAGCAGCA 982

Qy 661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCGGAA 720
Db 983 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCGGAA 1042

Qy 721 GAGTATCCAGAAAGTGCAAAGGTTTATGAAGAAACCATATGTCGATCCCCGTG 780
Db 1043 GAGTATCCAGAAAGTGCAAAGGTTTATGAAGAAACCATATGTCGATCCCCGTG 1102

Qy 781 AGCATGTCTGAACTGGAAAAGGTGGAAGTGGCTGCTCACCTGCTCTCAGTTTTAATTAAC 840
Db 1103 AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTCTCAGTTTTAATTAAC 1162

Qy 841 AAGAAGGTAGACTCCTGA 858
Db 1163 AAGAAGGTAGACTCCTGA 1180

```
Db   728 GATGAAAATGCAACTTTAGATGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTG 787
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 847
Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 907
Qy   541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGACAGAAGGCCCTTAAGATC 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGACAGAAGGCCCTTAAGATC 967
Qy   601 ATGCAACAGATGAGTGACCACCGTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 ATGCAACAGATGAGTGACCACCGTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 1027
Qy   661 AACTGTATATATCTAAATATCCCACAACAAAGGCACGTCTTGCTGCGACCCCCGAA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028 AACTGTATATATCTAAATATCCCACAACAAAGGCACGTCTTGCTGCGACCCCCGAA 1087
Qy   721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCGTG 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCGTG 1147
Qy   781 AGCATGTCTGAACTGGAAAAGGTGGATGGGCCTGCTCACCTGCTCCAGTTTTAATTAAC 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148 AGCATGTCTGAACTGGAAAAGGTGGATGGGCCTGCTCACCTGCTCCAGTTTTAATTAAC 1207
Qy   841 AAGAAGGTAGACTCCTGA 858
         ||||||||||||||||||
Db  1208 AAGAAGGTAGACTCCTGA 1225
```

RESULT 3
US-10-205-219-83
; Sequence 83, Application US/10205219
; Publication No. US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: N-G, N-G dimethylarginine dimethylaminohydrolase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2411, 2412
; OTHER INFORMATION: n is a or g or c or t
US-10-205-219-83

Query Match   83.4%;   Score 715.6;   DB 15;   Length 3008;
Best Local Similarity   89.6%;   Pred. No. 1.6e-195;
Matches 769;   Conservative   0;   Mismatches 89;   Indels   0;   Gaps   0;

```
Qy    1 ATGGCCGGCCTCGGCCACCCTCCGCCTTCGGCCGCGGCCACCCACCGCCGTCGGCCTGGGGCG 60
        |||||||||||| |||||||||| ||||||||||||||||||||||||||||||||| || ||
Db  432 ATGGCCGGCCTCAGCCACCCTCGGTCTTCGGCCGCGGCCACCCACCGCCGTCGGCCGTGCGGGCT 491
```

RESULT 4
US-10-252-157-328
; Sequence 328, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program

```
Qy   61 CTACCCGAGTGCTCTGCCAGCACGCACCGGCTGAGAACGGCCAAGGCGCCGAGGAGGTGGACGTC 120
        ||||||||| ||| |||||| |||| ||| |||||| |||||| ||| ||   ||    ||  | |
Db  492 CCGCCCGAGTCCCTGTGCCGCCACGGCGCTGAGGGCGCTGAGGCCGCCTCCCGGGCGAAGAGGTGGATTC 551
Qy  121 GCCCGCGCGGAACGGCAGCACCAGCTCTAGTGGCGCGTCAGTGGGCGTGCTGCGGCGGTGCGCAAGCTGGGGCTG 180
        ||      ||||| |||||||||||||| || ||||||   ||||||||||||||||||||||||||||||| |||
Db  552 GCTCGGCCTGAGCGCCAGCACCAGCTCTACGGGCGTGCTGCGGCGGTGCGCAAGCTGGGGCGTG 611
Qy  181 CAGGTGGTGGAGCTGCGGCCGCGGCCGCCGACGAGAGCGCCTTCCGGACGCCGGCCCGTCGTGCGGAGACGTG 240
        ||||||||| ||||||||||||||||||||||||||||||||||||||||  ||  ||||||||||||| ||||
Db  612 CAGGTGGTGCAGCTGCGGCCGCGGCCGCCGACGAGAGCGCCTTCCGGACGTCGTGTTCGTGCGGAGGACGTG 671
Qy  241 GCCGTGCGTGTGCGGAGGAGACGGCCCTCATCACCCGACACCGGCCGCGGGCCGGAGGAAG 300
        ||||| | |||||||||||||||||||| ||||||||| || ||    ||||||||||||
Db  672 GCCGTGTGTGTGCGGAGGAGACGGCCCCTCATCACCCGCCGCCGGGGGCCGGAGGAAG 731
Qy  301 GAGGTTGACATGATGAAGAGAGCATTAGAAACTTCAGCTCAATATAGTAGAGATGAAA 360
        ||||||||||||||||| ||||  |||| |||||||||||||| ||||||||||||||
Db  732 GAGGTTGACATGATGAAAGAGAGGCTTTGGAAAAAACTTCAGCTCAACATAGTAGAGATGAAA 791
Qy  361 GATGAAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 420
        ||||||||||||||||||||| | ||| ||     ||||||  |||||| ||||||||
Db  792 GATGAAAATGCAACTTTAGATGGTGGGGACGTTCAGATGGTGCTGAAATCTTTTAAGGAC 851
Qy  421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
        ||||||||||||||||||||||||||||||||||   ||||||||||| ||||| |||||
Db  852 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGTGGCTGGATCTTGGCTGTGATACTTTCAAGGAC 911
Qy  481 TATGCAGTCTCCACAGTGCCAGTGGTGCGCCAGTGGTTGCATTTGAAGAGTTTGCATTTGTG 540
        || |||||| ||||||| |||||||  | | ||   |||||||||| |||||| |||||| |
Db  912 TACGCAGTTTCCACAGTCCCAGTGGCCGATTCTTTGCATTTGAAAGAGTTTTGCATTTGCAGCATG 971
Qy  541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600
        |||||||| ||||||||||||||| |||||||||||||||| |||||||||||| |||||
Db  972 GCTGGGCCCCAACCTGATCGCAATAGGGTCCAGTGAATCTGCGCAGAAGGCCCTCAAGATC 1031
Qy  601 ATGCAACAGATGAGTGACCACCGTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660
        ||||||||||||| ||||||||| |||||||||||||||| ||||||||||||||||||
Db 1032 ATGCAACAGATGAGTGACCACCGTATGACAAAGGCTCACTGTACCGGACGACGCACATGGGCCGCC 1091
Qy  661 AACTGTATATATCTAAATATCCCACAACAAAGGCACGTCTTGCTGCGACCCCCGAA 720
        ||||||||||| |||||||| ||||   ||||| |||||| |||||||||||||||
Db 1092 AACTGTATATATTTAATATCCCAGCAAAAGGGGCAAAAGGCAAAAGGGAAACCCCAGAA 1151
Qy  721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCGTG 780
        |||||||||| |||||||||  |||||||||  || |||||| ||||| ||| ||| |||
Db 1152 GAGTACCCAGAAAGGCCAAAGTTTATGAGAAGCTCAAGGAGCTCAAGGACCATCCTACTGATCCCGTG 1211
Qy  781 AGCATGTCTGAACTGGAAAAGGTGGATGGGCCTGCTCACCTGCTCCAGTTTTAATTAAC 840
        ||||||| |||| |||||||||| ||||| ||||||||||| || |||||||||||| |
Db 1212 AGCAATTCTGAGATGGAAAAGGTGGACGGGCTTGCTCACCTGCTGCTCACCTGCTCCGTTTTTATTAAC 1271
Qy  841 AAGAAGGTAGACTCCTGA 858
        |||||| ||||| |||||
Db 1272 AAGAAGACAGACTCTTGA 1289
```

```
; SEQ ID NO 328
; LENGTH: 3778
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID ID NO. US20030190640A1 198087.1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1083-1106
; OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-328

Query Match          82.2%;  Score 705.4;  DB 15;  Length 3778;
Best Local Similarity  99.7%;  Pred. No. 1.6e-192;
Matches  717;  Conservative   0;  Mismatches   1;  Indels   1;  Gaps   1;

Qy     141 CCAGCTCTACGTGGGCGTGCTTGGGCAGCAGCTGCGGGCTGGGCGTGGTGGTGTGAGCTGCCGGC 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCAGCTCTACGTGGGGCGTGCTCGGCGCAGCAGCTGCGGGCTGGGCGTGGGCGTGGTGTGAGCTGCCGGC 60

Qy     201 CGACGAGAGCCTTCCGACTCGTCTTCGTGGAGGACGTGCCGTGTGTGCGAGGAGAC 260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CGACGAGAGCCTTCCGACTCGTCTTCGTGGAGGACGTGCCGTGGTGTGCGAGGAGAC 120

Qy     261 GGCCCTCATCACCCGACCCCGGGGCCGCCGGAGCCGGAGGAGGTTGACATGATGAAAGA 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGCCCTCATCACCCGACCCCGCCGGGGCCGCCGGAGCCGGAGGAGGTTGACATGATGAAAGA 180

Qy     321 AGCATTAGAAAAACTTCAGCTCAATATAGTAGATGAGATGAAAGATGAAAATGCAACTTTAGA 380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AGCATTAGAAAAACTTCAGCTCAATATAGTAGATGAGATGAAAGATGAAAATGCAACTTTAGA 240

Qy     381 TGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTGGGCCTTTCAAAAGGACAAA 440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTGGGCCTTTCAAAAGGACAAA 300

Qy     441 T-CAACGAGGTGCTGAAATCTTGGCTGTGATACTTTAAGGACTATGCAGTCTCCACAGTGC 499
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TCCAACGAGGTGCTGAAATCTTGGCTGTGATACTTTAAGGACTATGCAGTCTCCACAGTGC 360

Qy     500 CAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGCTGGGCCTAACCTGATCG 559
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGCTGGGCCTAACCTGATCG 420

Qy     560 CAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTGAGATCATGCAACAGATGAGTGACC 619
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTGAGATCATGCAACAGATGAGTGACC 480

Qy     620 ACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCATCTAAATA 679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCATCTAAATA 540

Qy     680 TCCCCAACAAAGGGCACAGTCTTGTGCTGCGACCCGAACCCCGGAAGAGTATCCAGAAAGTGCAA 739
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCCCCAACAAAGGGCACAGTCTTGTGCTGCGACCGCAGCCGGAAGAGTATCCAGAAAGTGCAA 600

Qy     740 AGGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCGTGAGCATGTCTGAACTGGAAA 799
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AGGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCGTGAGCATGTCTGAACTGGAAA 660

Qy     800 AGTGGATGATGGGCGTGCTCACCTGCTCACCTGCTGCTCAGTTTAATTAACAGAGGTAGACTCCTGA 858
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AGTGGATGGGGCGTGCTCACCTGCTCACCTGCTGCTCAGTTTAATTAACAGAGAAGTAGACTCCTGA 719
```

```
RESULT 5
US-09-925-300-512
; Sequence 512, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
```

```
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 512
; LENGTH: 3651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1283)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3641)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3650)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-512

Query Match          73.5%;  Score 630.4;  DB 9;  Length 3651;
Best Local Similarity  99.7%;  Pred. No. 6.7e-171;
Matches  642;  Conservative   0;  Mismatches   1;  Indels   1;  Gaps   1;

Qy     215 CGGACTGCGTCTTCGTGGAGGACGTGCCGTGGTGTGCGAGGAGACGGGCCCTCATCACCC 274
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 CGGACTGCGTCTTCGTGGAGGACGTGCCGTGGTGTGCGAGGAGACGGGCCCTCATCACCC 61

Qy     275 GACCCGGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAAC 334
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 GACCCGGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAAC 121

Qy     335 TTCAGCTCAATATAGTAGAGATGAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTT 394
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 TTCAGCTCAATATAGTAGAGATGAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTT 181

Qy     395 TATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTG 454
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 TATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTG 241

Qy     455 AAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGGT 514
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 AAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGGT 301

Qy     515 TGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTG 574
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 TGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTG 361

Qy     575 AATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCGCTACGACAAAC 634
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 AATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCGCTACGACAAAC 421

Qy     635 TCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCAACAAAGGGC 694
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 TCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCAACAAAGGGC 481

Qy     695 ACGTCTTGCTGCACCGAACCCGGAAGAGTATCCAGAAAGTGCAAAGGTTTATGAGAAAC 754
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 ACGTCTTGCTGCACCGAACCCGGAAGAGTATCCAGAAAGTGCAAAGGTTTATGAGAAAC 541

Qy     755 TGAAGGACCCATATGCTGATCCCGTGAGCATGTCTGAACTGGAAAAGTGGATGGGCTGC 814
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 TGAAGGACCCATATGCTGATCCCGTGAGCATGTCTGAACTGGAAAAGTGGATGGGCTGC 601

Qy     815 TCACCTGCTGCTCAGTTTTAATTAACAAGAGACTCCTGA 858
           ||||||||||||  ||||||||||||||||||||||||
Db     602 TCACCTGCTG-TCAGTTTTAATTAACAAGAAGTAGACTCCTGA 644
```

Query Match   21.6%;   Score 185.6; DB 17; Length 1397;
Best Local Similarity   57.4%;   Pred. No. 9.4e-43;
Matches 372; Conservative   0; Mismatches 274; Indels   2; Gaps   2;

```
Qy   29 TCGGCCGGGCCACCCACGCGTGGTGCGGGCCTCACCCGAGTCGCCTCTGCCAGCACGCGC 88
Db  357 TGGGCCGCTGCTCCCATGCCCGATCCGGGAGTCCCAGGAGACCTGGCGTCGGGGGAAG   416
Qy   89 TGAGAAGCGGCCAAGGGCGAGGAGGTGGAGCTGCCGCGCGGACGTCGCCGGGCACCAGCTCT 148
Db  417 GTGCCGGGCGGCTTCCCGCTCGGATCGGCCAAAGCTCAAAGGGGAGCAGAGGTGC  476
Qy  149 ACGTGGGCGGTGCTGGCAGCAAGTGGGCTGGGAGCTGGTGGAGCCGCCGACGAGA  208
Db  477 TGGGAGTAAACTGAGGCAACGACTGGGCTACGACTGCTGTGGAGGACGAGGGAGT  536
Qy  209 GCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGCTGTGTGAGGACGACGGCCCTCA 268
Db  537 CATTGCCGCTGGGACCGCTGCTTGGCGACACGGCCGTGATCCAAGGGACACGGCCCTAA  596
Qy  269 TCACCCGACCCGGGGCGGCCCCAGAGGAAGGAGGTTGACATGATGAAAGAAGCATTAG  328
Db  597 TCACGGCGCCCTGGAGCCCCGCTCGTAGGCCCAGAGGTCGCCGAAAGCCCTGC  656
Qy  329 AAAACTTCAGCTCAATATAGTAGAGGATGAAAGATGAAAATGCAACTTTAGATGCGGAG  388
Db  657 AAGACCTGGGGCTCCAATTGTGGAAATAGGAGAACGGAGCGCGACGCTGGATGCACTG  716
Qy  389 ATGTTTATTCACAGGACAGAGAATTTTTGTGGGGCCTTTCCAAAAGGACAAATCAACGAG  448
Db  717 ACGTTCTCTTCACCCGGGCCGGGAGTTTTTCGTAGGCCTCTCCAAATGGACCAATCACCGAG 776
```

Query Match   23.1%;   Score 197.8; DB 9; Length 770;
Best Local Similarity   61.6%;   Pred. No. 2.1e-46;
Matches 316; Conservative   0; Mismatches 197; Indels   0; Gaps   0;

```
Qy  110 AGGTGGACGTCGCCCGCGGAACGGCAGCACCAGCTCTACGTGGCGTGCTGGGCAGCA  169
Db  250 AGGTAGACCTGGCACGGCGCTCAACGTGAGAATGGAGTCTACTGTGTGCATCCTGAGACAGA 309
Qy  170 AGCTGGGGCTGCAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCG  229
Db  310 AACTTGGGCTGCAGGTGGTGGAGTTGCCCCCCAATGAGAACTGCCCCGGGGCCAATTGA  369
Qy  230 TGGAGACGTGCCGTGGTGTGCGAGGAGAGCAGGCCCTCATCACCCGACCCGGGGCGCCGA  289
Db  370 TAGGGGACACAGCTGTAGTGATAGCAGATACAGCCCTCATCACCCGTCCATGGATACCTG  429
Qy  290 GCCGGAGGAAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAG  349
Db  430 CACGAAGGAAAGAGACTGAAGGCCTGCAAAAAATCTTTGAGGAGCTGAAATTCCGAGTCT  489
Qy  350 TAGAGATGAAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTTTATTCACAGGCAGAG  409
Db  490 GCGAACTCAGTGATGAAAATGCCACTCTGGAACGTGCAAGTGATATACTTTTCACAGGTTCAG 549
Qy  410 AATTTTTTGTGGGGCCTTTCCAAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATA 469
Db  550 AGATTTTTGTAGGCTTGTCTAAATGACCAATCTTAGAGGTGCTGAAATGGTGGCAAAGA  609
Qy  470 CTTTAAGGACTATGCAGTCTCCAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTT  529
Db  610 CCTACCAGGATTATGCTGTTTCAACTGTTCCTGTGTCTGGGGACATGCCACTTTAAAAGCT  669
Qy  530 TCTGCAGCATGGCTGGGGCCTAACCTGATCGCCAATTGGGTTCAGTGAAATCTTGCACAGAAGG 589
Db  670 TTTGCAGTATGGCAGGGACCTGACACCCCTGGTCATAGGAAGCAGTGATACAGCGAGAAAGG 729
Qy  590 CCCTTAAGATCATGCAACAGATGAGTGACCACC 622
Db  730 CACTTGAAGACTATGACAACTGACCGATCATC 762
```

```
Qy  449 GTGCT-GAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCA 507
            ||  |||||||||||||||||||||| |||||||||||||||| |||||||
Db  777 GCGCTGGGAGATCGTGGCGGACACGTTCCGGGACTTCCGCCGTCTCCACTGTGCCAGTCTCG 836

Qy  508 GATGGGTTGCATTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGG 567
            |||||||||||  |||||||||||||||||||||||||||||||||||||
Db  837 GGTCCCTCCCACCTGCGCGGGTCTCTGCGCATGGGGGGACCTCGCACTGTTGTGGCAGGC 896

Qy  568 TCTAGTGAATCTGCACAGAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTAC 627
            |  ||||  |||||||||||||||||||
Db  897 AGCAGGGACGCGTGCCCAAAAGGCTGTCCGGGCAATGGCAGTCGTGCAGATCACCCATAT 956

Qy  628 GACAAACTCA-CTGTGCCTGATGACATAGCAGCAAACTGTATATATCT 674
            ||||||||||||||||
Db  957 GCCTCCCTGACCCCTCCAGATGACGACGCAGCTGCTGACTGTCTCTTTCT 1004
```

RESULT 8
US-09-783-590-5309
; Sequence 5309, Application US/09783590
; Patent No. US2002011085OA1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 5309
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (42)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (64)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (98)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (134)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (139)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (232)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (257)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (263)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (270)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (272)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (312)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (319)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (326)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-5309

Query Match   18.2%;   Score 156; DB 9; Length 332;
Best Local Similarity 84.3%; Pred. No. 1.6e-34;
Matches 253; Conservative 0; Mismatches 35; Indels 12; Gaps 7;

```
Qy  383 GCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTCCAAAAGGACAAATC 442
            ||||||||||||| ||||||||||||||||| |||||| ||| |||||||||||||
Db    7 GGGAAGATGTTTTATNCACAGGCAGAGAATTTTTTNTGGGCCTTTCCAAAAGGACAAANTC 66

Qy  443 AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTG-CCA 501
            ||||||||||||||||||||||||||||| ||||||||||||||||||||||||| |||
Db   67 AACGAGGTGCTGAAATCTTGGCTGATACTTTNAAGGACTATGCAGTCTCCACAGTGCCCA 126

Qy  502 GTGGCAG-ATGGGGTTGCATTGAAGAGTTTCTG-CAGCATGGCTGGGC---TAACCTGA 556
            ||||||| |||| |||||||| ||||||||||  |||||||| ||||||   |||||||
Db  127 GTGGCAGNATGGNTTGCATTTAAGAGTTTCTGCGCCAGCATGGTTGGGCCCTGAACCTGAA 186

Qy  557 TCGCAATTGGGTCTAGTGAA--TCTGCACAGAGAGGCCCTTAAGATCATGCAACAGATGAG 614
            |||||||||||| | ||||  ||| |||||| ||||||| ||||||| ||||||| ||||
Db  187 TCGCAATTGGGTGTGTGAAATCTGCCACAGAGGCCCCTTAAGATNATGCAACAAATGAG 246

Qy  615 TGA--CCACCGCTACGACAAACTCACTGTGCC--TGATGACATAGCAGCAAACTGTATAT 670
            |||  |||||| || |  |||| | ||||||  ||||:|||||||||||||  ||||||
Db  247 TGAACCACCGNTAAGGNCAAACTNANTGTGCCCTGAT:GACATAGCAGCAAATTGTATAT 306
```

RESULT 9
US-09-918-995-9117
; Sequence 9117, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9117
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-9117

Query Match   17.3%;   Score 148.4; DB 10; Length 472;
Best Local Similarity 61.3%; Pred. No. 2.9e-32;
Matches 239; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```
Qy  113 TGGACGTCGCCCGCGGCCGAACGGCACCAGCTCACGTGGGCGTGCTGGGCAGCAAGC 172
            |||| ||| ||||| ||  |  || |  ||||    | ||| |  ||   ||   ||
Db   82 TGGATCTGGCCAAAGCTCAAAGGCACTCAAAGGGCACACGGGGGGCTGGGAGGTAAACTCAGGCAACGAC 141

Qy  173 TGGGGCTGCAGGTGGTGTGGAGTCCGGGAGCTGCCGGCCGCGAGAGAGCGCTTCCGGACTCGTCTTCGTGG 232
```

```
Db   142                                                  201
         |||| |||| ||||||| || ||| |||||  ||||  ||||||||||||
Qy   142 TGGGGTTACAGCTGCTAGAACTGCCACCTGAGGAGGCATTGCCGCTGGGACCGCTGCTTG 201
Db   233 AGGACGTGGCCGTGGTGTGCGAGGAGGACGGCCTCATCACCGACCCGGCGGCCGCCGAGCC 292
             |||   ||| |||  | ||||  ||||||
Qy   202 GCGCACGCCCAGATCAAGGGACACGGCGGCCCCTAATCACGGCGCCCTGGAGCCCCGCTC 261
Qy   293 GGAGGAAGGAGGTTGACATGGATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAG 352
             ||||| ||| | |||| ||||
Db   262 GTAGGCCAGAGGTCGATGGAGTCCGCAAGACCTGGGCTCCGAATGTGG 321
Qy   353 AGATGAAAGATGAAAATGCAACTTAGATGGGGAGAGTGTTTATTCACAGGCAGAGAAT 412
             |||||||  ||
Db   322 AAATAGGAGACAGAACGGCGACGCTGACCGTTCTTTCACCGGCCGGGAGT 381
Qy   413 TTTTTGTGGGCCTTTCCAAAAGACAAATCAACGAGTGCTGAAATCTTGGCTGATACTT 472
             |||| |||||||| |||||  || ||||||| |||||||
Db   382 TTTTCGTAGGCCTTCCCAAATGACCAATCACCGAGACCTGAGATCGTGCGGACACGT 441
Qy   473 TTAAGGACTATGCAGTCTCCCAGTGCCAG 502
             |||| ||||||||||||||||
Db   442 TCCGGGACTTCGCGCGTCTACACTGTGCCAG 471

RESULT 10
US-10-450-826-8/c
; Sequence 8, Application US/10450826
; Publication No. US2004101818A1
; GENERAL INFORMATION:
;   APPLICANT: Ji, Darren
;   APPLICANT: Axelrod, Douglas W.
;   APPLICANT: Cook, Jonathon S.
;   APPLICANT: Jaiswal, Neelam
;   APPLICANT: Einstein, Richard
;   APPLICANT: Houghton, Adam
;   APPLICANT: Mertz, Lawrence
;   TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiati
;   FILE REFERENCE: 044921-5039-WO
;   CURRENT APPLICATION NUMBER: US/10/450,826
;   CURRENT FILING DATE: 2003-06-18
;   PRIOR APPLICATION NUMBER: US 60/255,882
;   PRIOR FILING DATE: 2000-12-18
;   PRIOR APPLICATION NUMBER: US 60/285,691
;   PRIOR FILING DATE: 2001-04-24
;   NUMBER OF SEQ ID NOS: 149
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 83946
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. AL078459
US-10-450-826-8

Query Match     17.2%;  Score 148;  DB 17;  Length 83946;
Best Local Similarity   100.0%;  Pred. No. 5.5e-31;
Matches 148;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    596 AGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAG 655
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  18612 AGATCATGCAACAGATGAGTGACCACCGTACGACAAACTCACTGTGCCTGATGACATAG 18553
Qy    656 CAGCAAACTGTATATATCTAAATATCCCAACAAAAGGCACGTCTTGCTGACCGAACCC 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  18552 CAGCAAACTGTATATATCTAAATATCCCAACAAAAGGCACGTCTTGCTGACCGAACCC 18493
Qy    716 CGGAAGAGTATCCAGAAAGTGCAAAGGT 743
          ||||||||||||||||||||||||||||
Db  18492 CGGAAGAGTATCCAGAAAGTGCAAAGGT 18465

RESULT 11
US-09-918-995-14871
```

```
; Sequence 14871, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;   APPLICANT: Hyseq, Inc.
;   TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;   TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;   FILE REFERENCE: 20411-756
;   CURRENT APPLICATION NUMBER: US/09/918,995
;   CURRENT FILING DATE: 2001-07-30
;   PRIOR APPLICATION NUMBER: US/09/235,076
;   PRIOR FILING DATE: 1999-01-20
;   NUMBER OF SEQ ID NOS: 38054
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14871
;   LENGTH: 452
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(452)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-14871

Query Match     17.0%;  Score 146;  DB 10;  Length 452;
Best Local Similarity   58.6%;  Pred. No. 1.4e-31;   Indels  0;  Gaps  0;
Matches 242;  Conservative  0;  Mismatches 171;

Qy    98 CCAAGGGCGAGGAGGTGGACGTGCGCCCGCGGCGGAACGGCAGCACCAGCTCTACGTGGGCG 157
         |||  | ||||||| ||   | | | |  |       ||||  |||| ||    ||||| |
Db    34 CAACTGGCGAGCAAANCNNNAAATNGNAAAGNNAAGGAGCACGGGGNCTGGGAGGNA 93
Qy   158 TGCTGGCAGCAAGTGGGGTGCAGGTGGTGGAGCTGCCGCCGACGAGAGCCTTCCGG 217
         || | ||  |||| ||| || || |||   | ||   |||| ||| |||| ||||
Db    94 AACTGTGGCCAACGACTGTTGCTACAGCTGCTAGAACTGCCACCTGAGGAGTCAITGCCGC 153
Qy   218 ACTGGCGTCTTCGTGAGGACGTGGCCGTGGTGTGCGAGGAGACGGCCCTCATCACCGAC 277
         |||||| | | || ||| |  || ||  | | ||| |  ||  |||  | |||| |
Db   154 TGGGACCGCTGCTTGGCGCACAGCGCCGTGATCCAAGGGACACGGGCCCTAATCACCGGC 213
Qy   278 CCGGGGCGCCGAGCCGGAAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTC 337
         |||   | | || || |||||| |||| | ||||| |   ||  |||| |
Db   214 CCTGAGCCCCGCTCGTAGGCCAGAGGTCGATGGAGTCGGCAAGACCCTGCAAGACCTGG 273
Qy   338 AGCTCAATATAGTAGTAGATGAAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTTAT 397
         |||| |||| ||| |||| | ||| | ||  | |  |  |||| | ||| ||  ||  |
Db   274 GGCTCCGAATTGTGGAAATAGGAGAGACGAGACGCGACGCTGGATGGCACTGACGTTCTCT 333
Qy   398 TCACAGGCAGAGAATTTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAA 457
         |||| |  |||||| | || ||| | |  |  ||||   |||| ||  |||||
Db   334 TCACCGGCCGGGAGTTTTTCGTAGGCCTCTCCAAATGACCAATCACCGAGGAGCTGAGA 393
Qy   458 TCTTGGCTGATACTTTTTAAGGACTATGCAGTCTCCAGTGCCAGAT 510
         || | |||  | || ||| | |||| ||| |||||| |||||||||
Db   394 TCGTGGCGGCACGTTCGCGGGACTTCGCGCGTCTCCACTGTGCCAGTCTCGGGT 446

RESULT 12
US-09-918-995-19254
; Sequence 19254, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;   APPLICANT: Hyseq, Inc.
;   TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;   TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;   FILE REFERENCE: 20411-756
;   CURRENT APPLICATION NUMBER: US/09/918,995
;   CURRENT FILING DATE: 2001-07-30
;   PRIOR APPLICATION NUMBER: US/09/235,076
;   PRIOR FILING DATE: 1999-01-20
;   NUMBER OF SEQ ID NOS: 38054
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19254
;   LENGTH: 403
```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-19254

Query Match   15.6%;   Score 134.2; DB 10; Length 403;
Best Local Similarity 59.6%;   Pred. No. 3.4e-28;
Matches 226; Conservative   0; Mismatches 153; Indels   0; Gaps   0;

Qy   121 CCCCGCCGGAACGGCAGCACCAGCTTCACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 180
               ||||||| |||||||||||||||||||
Db   24 GGCCAAGCTCAAAGGGACACGGGGTGCTGGGAGGTAAACTGAGGCAACGACTGGGGCTA   83

Qy   181 CAGGTGGTGGAGCTGCCGCGCCGACGAGAGCCTTCCCGACTGCGTCGTCGTCGAGGACGTG 240
               ||| |||||||||| |||||||||| |||||||
Db   84 CAGCTGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCGCCGCTGCTTGGCGACACG 143

Qy   241 GCCGTGGTGTCGGAGGACGGACCGCCCTCATCACCCGACACCCGGGCGCCGGAGGAAG 300
               ||| |||
Db   144 GCCGTGATCCAAGGGACACGGGACCTAATCACGCGGCCCTGAGCCCGCTCGTAGGCCA 203

Qy   301 GAGGTTGACATGACATGAAAGAAGAGCATTAGAAAACTTCAGCTCAATATAGTAGGATGAAA 360
               |||||||
Db   204 CAGGTGCGATGGAGTCCGCAAAGCCCTGCAAGACCCTGGGGCTTCGGAATTGTGGAAATAGGA 263

Qy   361 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 420
               ||||| ||||| |||||||||
Db   264 GACGAGAACGCGACGCGTGGATGCACTGACGTTCTGTCACCGGCGCGGGAGATTTTCAGA 323

Qy   421 GGCCTTTCCAAAAGGACAAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
               ||| |||||||||||||
Db   324 GGCCTCTCCTAATGGACCAATCACCGAGGAGCTGAGATCGTGGCCGGACACGTTCCGGGAC 383

Qy   481 TATGCAGTCTCCACAGTGC 499
               |||||| |||||
Db   384 TTCGCAGTCTTCCACTGTGC 402

RESULT 13
US-09-814-353-20553/c
; Sequence 20553, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
;   APPLICANT: Lee, John
;   APPLICANT: Thompson, Pamela
;   APPLICANT: Lillie, James
;   TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;   TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;   TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
;   FILE REFERENCE: MRI-006B
;   CURRENT APPLICATION NUMBER: US/09/814,353
;   CURRENT FILING DATE: 2001-03-21
;   PRIOR APPLICATION NUMBER: US 60/191,031
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: US 60/207,124
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 60/211,940
;   PRIOR FILING DATE: 2000-06-15
;   PRIOR APPLICATION NUMBER: US 60/216,820
;   PRIOR FILING DATE: 2000-07-07
;   PRIOR APPLICATION NUMBER: US 60/220,661
;   PRIOR FILING DATE: 2000-07-25
;   PRIOR APPLICATION NUMBER: US 60/257,672
;   PRIOR FILING DATE: 2000-12-21
;   NUMBER OF SEQ ID NOS: 22037
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20553
;   LENGTH: 3189
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1, 2, 3189
;   OTHER INFORMATION: n = A,T,C or G

US-09-814-353-20553

Query Match   13.9%;   Score 119.4; DB 10; Length 3189;
Best Local Similarity 99.2%;   Pred. No. 1.8e-23;
Matches 120; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   738 AAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGA 797
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3187 AAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGA 3128

Qy   798 AAAGGTGGATGGGGTGCTCACCTGCTCCAGTTTAATTAACAAGAAGGTAGACTCCTG 857
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3127 AAAGGTGGATGGGGTGCTCACCTGCTCCAGTTTTAATTAACAAGAAGTAGACTCCTG 3068

Qy   858 A 858
          |
Db 3067 A 3067

RESULT 14
US-09-736-457-404/c
; Sequence 404, Application US/09736457
; Patent No. US20020168837A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   APPLICANT: Fan, Liqun
;   APPLICANT: Wang, Aijun
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C15
;   CURRENT APPLICATION NUMBER: US/09/736,457
;   CURRENT FILING DATE: 2000-12-13
;   NUMBER OF SEQ ID NOS: 1864
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
;   LENGTH: 239
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-736-457-404

Query Match   9.5%;   Score 81.2; DB 9; Length 239;
Best Local Similarity 58.8%;   Pred. No. 5.1e-13;
Matches 140; Conservative   0; Mismatches 98; Indels   0; Gaps   0;

Qy   185 TGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCGTCTTCGTGAGGACGTGGCCG 244
               ||||||||||||| |||||||| |||||||| |||||
Db   238 TGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCGCTGCTTGGCGACACGGCCG 179

Qy   245 TGGTGTGCGAGGAGACGGCGCCCTCATCACCCGACCCGGGCGGGAGGAAGGAGG 304
               ||||| ||| |||||| |||||||| ||||||
Db   178 TGATCCAAGGGGACACGGGGCCCTAATCACGCGGCCCTGAGCCCGCTCGTAGGCCAGAGG 119

Qy   305 TTGACATGATGAAAGAAGCATTAGAAAACTTCAGCTCAATATAGTAGGATGAGAAAGATG 364
               ||| ||||| |||| |||||||
Db   118 TCGATGGAGTCCGCAAAGCCCTGCAAGACCCTGGGGCTCCGAATTGTGGAAATAGGAGACG   59

Qy   365 AAAATGCAACTTTAGATGGGGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGG 422
               |||| |||| ||||||
Db   58 AGAACCGCGACCGCTGGATGGCACTGACGTTCTTCACCGGCGCGGGAGTTTTTCGTAGG 1

RESULT 15
US-09-902-941-404/c
; Sequence 404, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:

;   APPLICANT: Henderson, Robert A.
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Watanabe, Yoshihiro
;   APPLICANT: Johnson, Jeffrey C.
;   APPLICANT: Retter, Marc W.
;   APPLICANT: Marnerakis, Margarita
;   APPLICANT: Carter, Darrick
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Vedvick, Thomas S.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: McNabb, Andria
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C17
;   CURRENT APPLICATION NUMBER: US/09/902,941
;   CURRENT FILING DATE: 2001-07-10
;   NUMBER OF SEQ ID NOS: 2002
;   SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 404
;   LENGTH: 239
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-902-941-404

Query Match   9.5%; Score 81.2; DB 9; Length 239;
Best Local Similarity   58.8%; Pred. No. 5.1e-13;
Matches 140; Conservative   0; Mismatches   98; Indels   0; Gaps   0;

Qy 185 TGGTGGAGCTGCCGGCCGAGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCG 244
            ||||||| ||| || ||||| |||| | ||||||||||| ||||||| |||||
Db 238 TGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCGCTGCTTGGCGACACCGCCG 179

Qy 245 TGGTGTGCGAGGAGGACGGCCTCATCACCCGACCCGGGGCGCCGAGCCGGAAGGAGG 304
            ||||||| ||||| |||||| || ||| |||||||| || ||| |||||||
Db 178 TGATCCAAGGGACACGGCCCTAATCACGCGGCCCTGGAGCCCCGCTCGTAGGCCAGAGG 119

Qy 305 TTCACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATG 364
            |||| || ||||||| | | || | ||| || || | |||
Db 118 TCCATGGAGTCCGCAAAGCCCTGCAAGACCCTGGGGCTTCCGAATTGTGGAAATAGGAGACG 59

Qy 365 AAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTGGG 422
            || || | ||| || ||| | ||||| ||| ||| ||||| ||||||
Db   58 AGAACGCCGACGCTCGATGCCACTGACGTTCTCTTCACCGGCCCGGGAGTTTTTCGTAGG 1

Search completed: June 7, 2004, 21:24:56
Job time : 430 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   June  7, 2004, 20:25:50 ; Search time 404 Seconds
                                    (without alignments)
                         9022.163 Million cell updates/sec

Title:           US-09-889-733B-1
Perfect score:   858
Sequence:        1 atggccggctcggccaccc.......acaagaggtagactcctga 858

Scoring table:   OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched:   3373863 seqs, 2124099041 residues

Word size :   0

Total number of hits satisfying chosen parameters:   6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
             1: geneseqn1980s:*
             2: geneseqn1990s:*
             3: geneseqn2000s:*
             4: geneseqn2001as:*
             5: geneseqn2001bs:*
             6: geneseqn2002s:*
             7: geneseqn2003as:*
             8: geneseqn2003bs:*
             9: geneseqn2003cs:*
            10: geneeqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

```
Result            %     Query
 No.   Score  Match Length DB    ID         Description
------ -----  ----- ------ --  -------       -----------
   1    858   100.0   858   3  AAA47655    Aaa47655 Dimethyla
   2    772    90.0   858   9  ADE71465    Ade71465 Human Thr
   3    772    90.0  1633   4  AAH02907    Aah02907 Human she
   4    772    90.0  1633   6  ABQ88102    Abq88102 Human ost
   5    772    90.0  3932   7  ABZ35939    Abz35939 Human sec
   6    772    90.0  4208   6  AAL41494    Aal41494 Drug meta
   7    721    84.0   858   9  ADE71463    Ade71463 Human Met
   8    610    71.1  3651   3  AAF16077    Aaf16077 Human pro
   9    585    68.2  3778   9  ADE53981    Ade53981 Human pro
c 10    148    17.2  83946   9  ABQ88101    Abq88101 Human ost
  11    114    13.3  6373   5  ABA14642    Aba14642 Human ner
  12    114    13.3  6381   5  ABA14641    Aba14641 Human ner
  13     85     9.9   420   3  AAC10265    Aac10265 Human sec
  14     60     7.0    60   6  ABN36282    Abn36282 Human spl
  15     54     6.3  3008   9  ADB53014    Adb53014 Primary r
  16     54     6.3  3008   6  ADB79843    Adb79843 Rat N-G,N
  17     23     2.7  2523   7  ACC69598    Acc69598 Rat sweet
  18     23     2.7  2771   3  AAZ50743    Aaz50743 Rat senso
  19     22     2.6   862   6  ABQ47888    Abq47888 Oligonucl
c 20     22     2.6   862   6  ABQ47889    Abq47889 Oligonucl
c 21     22     2.6   898   6  ABQ41319    Abq41319 Oligonucl
c 22     22     2.6   898   6  ABQ41318    Abq41318 Oligonucl
  23     21     2.4    21   3  AAA47661    Aaa47661 Expressed
  24     21     2.4    21   3  AAA47663    Aaa47663 Expressed
  25     21     2.4    21   3  AAA47664    Aaa47664 Expressed
  26     20     2.3   444   3  ABN89945    Abn89945 Mouse clo
  27     20     2.3  21470   8  ABK42270    Abk42270 Genomic s
  28     20     2.3  21470   6  ADB60426    Adb60426 Connectiv
c 29     20     2.3  110000  6  ABX08336    Abx08336_00 Human pho
  30     19     2.2   314   2  AAQ60489    Aaq60489 Human bra
  31     19     2.2   368   4  ABK42398    Abk42398 Genomic s
c 32     19     2.2   368   4  AAK68647    Aak68647 Human imm
c 33     19     2.2   368   8  ADB60554    Adb60554 Connectiv
c 34     19     2.2   573   4  AAL46258    Aal46258 Murine gr
  35     19     2.2   573   4  AAL46259    Aal46259 Murine gr
  36     19     2.2   718   4  AAL46254    Aal46254 Murine gr
  37     19     2.2   718   4  AAL46255    Aal46255 Murine gr
  38     19     2.2   758   2  AAT62819    Aat62819 Murine gr
  39     19     2.2   777   4  AAL46224    Aal46224 Murine gr
c 40     19     2.2   954   7  ACA21957    Aca21957 Prokaryot
  41     19     2.2  2481   6  ABT10749    Abt10749 Human bre
  42     19     2.2  4412   9  ADD24915    Add24915 DNA encod
  43     19     2.2  7084   7  ACC43645    Acc43645 Nucleotid
  44     19     2.2  7263   4  AAI58179    Aai58179 Human pol
  45     19     2.2  7263   8  ADB48145    Adb48145 Novel hum
```

ALIGNMENTS

```
RESULT 1
AAA47655
ID   AAA47655 standard; cDNA; 858 BP.
XX
AC   AAA47655;
XX
DT   08-NOV-2000  (first entry)
XX
DE   Dimethylarginine dimethylaminohydrolase (DDAH1) coding sequence.
XX
KW   Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH2;
KW   arginine deaminase; hyperlipidemia; renal failure; hypertension;
KW   restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;
KW   ischemia reperfusion injury; septic shock; multi organ failure;
KW   arthritis; skin disorders; inflammatory cardiac disease; migraine;
KW   infection; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..858
FT                   /*tag= a
FT                   /product= "DDAH1"
XX
PN   WO20004888-A2.
XX
PD   03-AUG-2000.
XX
DR   WPI; 2000-543392/49.
DR   P-PSDB; AAB01477.
XX
PF   26-JAN-2000; 2000WO-GB000226.
XX
PR   26-JAN-1999;   99GB-00001705.
PR   04-JUN-1999;   99GB-00013066.
XX
PA   (UNLO ) UNIV COLLEGE LONDON.
XX
PI   Vallance PJT, Leiper JM, Whitley GSJ,  Charles IG;
XX
PT   Novel methylarginase polypeptides and polynucleotides, used to identify
PT   modulators of them, which are used in the treatment of e.g. cancer,
PT   hypertension, and bacterial infections.
XX
PS   Claim 1; Page 55-56; 68pp; English.
XX
```

CC   Nucleotides encoding methylarginase polypeptides, vectors comprising
CC   these nucleotides and the polypeptides themselves can be used in
CC   medicaments for the treatment of hyperlipidemia, renal failure,
CC   hypertension, restenosis after angioplasty, atherosclerosis,
CC   complications of heart failure, schizophrenia, multiple sclerosis or
CC   cancer. Modulators of the enzyme can be used in medicaments for the
CC   treatment of ischemia-reperfusion injury of the brain or heart, cancer,
CC   lethal hypertension in severe inflammatory conditions such as septic
CC   shock or multi-organ failure, or local and systemic inflammatory
CC   disorders including arthritis, skin disorders, inflammatory cardiac
CC   disease, migraine, or microbial or bacterial infection
XX
SQ   Sequence 858 BP; 217 A; 214 C; 256 G; 171 T; 0 U; 0 Other;

Query Match   100.0%; Score 858; DB 3; Length 858;
Best Local Similarity   100.0%;
Matches 858; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGGCCGGCCTCGGCCACCCCTCCGCCTTCGGCCGGCCACCCAGCCGCCTGGTGCGGCGGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGCCGGCCTCGGCCACCCCTCCGCCTTCGGCCGGCCACCCAGCCGCCTGGTGCGGCGGG 60

Qy   61 CTACCCGAGTCGTCTGCCAGCACCGCGCTGAGAAGCGCCAAGGGCGGAGGAGGTGGACGTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTACCCGAGTCGTCTGCCAGCACCGCGCTGAGAAGCGCCAAGGGCGGAGGAGGTGGACGTC 120

Qy   121 GCCCGCGGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGGCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCCCGCGGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGGCTG 180

Qy   181 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCGGACTGCGCTTCGGAGGACGTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCGGACTGCGTTCGGAGGACGTG 240

Qy   241 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGGGCCGCCGAGCCCGGAGGAAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGGGCCGCCGAGCCCGGAGGAAG 300

Qy   301 GAGGTTGACATGATGAAAGAAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GAGGTTGACATGATGAAAGAAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360

Qy   361 GATGAAAATGCAACTTTAGATGGCCGAGATGTTTATTCACAGGCAGAATTTTTTGTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GATGAAAATGCAACTTTAGATGGCCGAGATGTTTATTCACAGGCAGAATTTTTTGTG 420

Qy   421 GGCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 480

Qy   481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TATGCAGTCTCCACAGTGCCAGTGGGTTGCATTGAAGAGTTTCTGCAGCATG 540

Qy   541 GCTGGGCCTAACCTGATCGCAATTGGGTCTCAGTGAATCTGCACAGAAGGCCCCTTAAGATC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCTGGGCCTAACCTGATCGCAATTGGGTCTCAGTGAATCTGCACAGAAGGCCCCTTAAGATC 600

Qy   601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGCCTGATGACATAGCAGCA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGCCTGATGACATAGCAGCA 660

Qy   661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCCGGAACCCCGGAA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGCGAACCCCGGAA 720

Qy   721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACCAATAGCTGATCCCCGTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCAATAGCTGATCCCCGTG 780

Qy   781 AGCATGTCTGAACTGGAAAAGGTGATGAAAGTGGAGATGGGGCTCACCTGGCTCACCGTGCTTAATTAAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGCATGTCTGAACTGGAAAAGGTGATGGGGCTCACCTGGCTCACCGTGCTTAATTAAC 840

Qy   841 AAGAAGGTAGACTCCTGA 858
              ||||||||||||||||||
Db   841 AAGAAGGTAGACTCCTGA 858

RESULT 2
ADE71465
ID   ADE71465 standard; DNA; 858 BP.
XX
AC   ADE71465;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human Thr87 wild-type dimethylarginine dimethylaminohydrolase-1 DNA.
XX
KW   Human; dimethylarginine dimethyl-aminohydrolase-1; DDAH 1; chromosome 1;
KW   single nucleotide polymorphism; SNP; ADMA;
KW   asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
KW   coronary heart disease; cerebrovascular disease; hypertension; diabetes;
KW   susceptibility; genotyping; therapeutic targeting; antidiabetic;
KW   cardiant; cerebroprotective; hypotensive; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..858
FT   /*tag= a
FT   /product= "Wild-type (Thr87) DDAH 1"
FT   variation   /*tag= b
FT   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   /note= "Results in Thr (C variant) or Met (T variant) at
FT   position 87 of the protein"
XX
PN   WO200308963B-A1.
XX
PD   30-OCT-2003.
XX
PF   11-APR-2003; 2003WO-FI000274.
XX
PR   19-APR-2002; 2002US-00125456.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Valkonen V,   Salonen JT,   Pirskanen M,   Tuomainen T,   Laakso J;
PI   Laaksonen R;
XX
DR   WPI; 2001-854121/79.
DR   P-PSDB; ADE71466.
XX
PT   New nucleic acid encoding a variant dimethylarginine
PT   dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
PT   and its vascular complications, e.g. coronary or cerebrovascular disease
PT   or hypertension.
XX
PS   Disclosure; SEQ ID NO 3; 37pp; English.
XX
CC   The invention relates to a variant dimethylarginine
CC   dimethylaminohydrolase-1 (DDAH 1; ADE71464) and a nucleic acid encoding
CC   it (ADE71463). DDAHs regulate the metabolism of ADMA (asymmetrical
CC   dimethyl-arginine, a naturally occurring nitric oxide synthase inhibitor)
CC   and nitric oxide, which are relevant in disorders such as cardiovascular
CC   disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
CC   has been implicated in susceptibility to diseases such as familial
CC   combined hyperlipidaemia, premature coronary artery disease (CAD), non
CC   insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
CC   The variant DDAH 1 of the invention has Met at position 87, rather than
CC   Thr, which is caused by a C to T change in exon 1 at a position
CC   corresponding to base 260 in ADE71463 and ADE71465. Individuals hetero-
CC   or homozygous for the Thr87Met allele have an increased risk of
CC   developing cardiovascular diseases and diabetes. The invention also
CC   relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a

CC   susceptibility to a cardiovascular disease and diabetes in an individual
CC   by determining the DDAH 1 genotype; a method of treating diabetes or
CC   vascular complications of diabetes by enhancing nitric oxide
CC   availability, production or concentration; a method of targeting
CC   treatment of cardiovascular disease and diabetes in a hypertensive
CC   patient by determining the DDAH 1 genotype and treating them with a drug
CC   that affects nitric oxide availability, production or metabolism; and a
CC   transgenic animal which carries a human variant DDAH 1 nucleic acid
CC   sequence. The nucleic acid molecules and polypeptides are useful for
CC   treating cardiovascular disease (e.g., coronary heart disease,
CC   cerebrovascular disease, and hypertension), and diabetes and its vascular
CC   complications. The methods are useful for determining whether a patient
CC   will benefit from treatment with a drug which affects nitric oxide
CC   availability, production or metabolism; a drug which reduces ADMA
CC   availability or concentration; orn an agent which elevates DDAH
CC   availability or concentration (such as DDAH agonist). The methods are
CC   also useful for determining whether a patient will be at risk of adverse
CC   effects if DDAH antagonists are administered. The present sequence
CC   represents a nucleic acid encoding the "wild-type" (Thr8") DDAH 1.
XX
SQ   Sequence 858 BP; 218 A; 213 C; 258 G; 169 T; 0 U; 0 Other;

Query Match   90.0%; Score 772; DB 9; Length 858;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 822; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   23 CGGCCTTCGGCCGGGCCACCCAGCGCGTGGTGCGGGGCGGCTACCCGAGTCGCTCTGCCAGC 82
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   23 CGGCCTTCGGCCGGGCCACCCAGCGCGTGGTGCGGGGCGGCTACCCGAGTCGCTCTGCCAGC 82

Qy   83 ACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGACGTGCCCCGCGGAACGGCAGCACC 142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83 ACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGACGTGCCCCGCGGAACGGCAGCACC 142

Qy   143 AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGGTGGAGCTGCCGGCCG 202
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   143 AGTCCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGGTGGAGCTGCCGGCCG 202

Qy   203 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAAGCAGTGCCGTTGGTGTGCGAGGAGACGG 262
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAAGCAGTGCCGTTGGTGTGCGAGGAGACGG 262

Qy   263 CCCTCATCACCCGACACCCGGGGCGCGGAGCCGGAGGAAGGAGGTTGACATGATGAAAGAAG 322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263 CCCTCATCACCCGACACCCGGGGCGCGGAGCCGGAGGAAGGAGGTTGACATGATGAAAGAAG 322

Qy   323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAATGCAACTTTAGATG 382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAATGCAACTTTAGATG 382

Qy   383 GCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC 442
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   383 GCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC 442

Qy   443 AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACATGCAGTCTCCACAGTGCCAG 502
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   443 AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACATGCAGTCTCCACAGTGCCAG 502

Qy   503 TGGCAGATGGGTTGCATTGAAAGAGTTTCTGCAGCATGCAGTCGGCCTAACCTGATCGCAA 562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   503 TGGCAGATGGGTTGCATTGAAAGAGTTTCTGCAGCATGCAGTCGGCCTAACCTGATCGCAA 562

Qy   563 TTGGGTCTAGTGAATCTGCACAGAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCACC 622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   563 TTGGGTCTAGTGAATCTGCACAGAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCACC 622

Qy   623 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 682
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   623 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 682

Qy   683 CCAACAAAGGGCACGTCTTGCTCACCGAACCCCGGAGAATCCGAGAGTATCCAGAAAGTGCAAAGG 742
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   683 CCAACAAAGGGCACGTCTTGCTCACCGAACCCCGGAGAATCCGAGAGTATCCAGAAAGTGCAAAGG 742

Qy   743 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTGCTGAACTGGAAAAGG 802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   743 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTGCTGAACTGGAAAAGG 802

Qy   803 TGGATGGGCTGCTCACCTGCTCAGTTTAATTAACAAGAA 845
             |||||||||||||||||||||||||||||||||||||||
Db   803 TGGATGGGCTGCTCACCTGCTCAGTTTAATTAACAAGAA 845

RESULT 3
AAH02907
ID   AAH02907 standard; DNA; 1633 BP.
XX
AC   AAH02907;
XX
DT   15-JUN-2001 (first entry)
XX
DE   Human shear stress-response coding sequence SEQ ID NO: 67.
XX
KW   Human; shear stress-response protein; vascular disease; arteriosclerosis;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200125427-A1.
XX
PD   12-APR-2001.
XX
PF   02-OCT-2000; 2000WO-JP006840.
XX
PR   01-OCT-1999;   99JP-00280976.
XX
PA   (KYOW) KYOWA HAKKO KOGYO KK.
PA   (NOJI/) NOJIMA H.
XX
PI   Nojima H, Yoshisue H, Obayashi M, Ota T,   Kawabata A,   Sakurada K;
PI   Kuga T, Sekine S, Nakamura Y, Sugano S;
XX
DR   WPI; 2001-266308/27.
DR   P-PSDB; AAB90784.
XX
PT   DNA sequences, proteins encoded by them and antibodies against them
PT   useful in diagnosis and treatment of vascular disease caused by
PT   arteriosclerosis.
XX
PS   Claim 20; Page 404-407; 678pp; Japanese.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human shear stress response proteins. These are useful in the
CC   diagnosis, treatment and screening of vascular diseases caused by
CC   arteriosclerosis, including heart failure, post-PTCA restenosis and
CC   hypertension
XX
SQ   Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Query Match   90.0%; Score 772; DB 4; Length 1633;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 822; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   23 CGGCCTTCGGCCGGGCCACCCAGCGCGTGGTGCGGGGCGGCTACCCGAGTCGCTCTGCCAGC 82
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   345 CGGCCTTCGGCCGGGCCACCCAGCGCGTGGTGCGGGGCGGCTACCCGAGTCGCTCTGCCAGC 404

Qy   83 ACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTGCCCCGCGGAACGGCAGCACC 142
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   405 ACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTGCCCCGCGGAACGGCAGCACC 464

Qy   143 AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGGTGGAGCTGCCGGCCG 202
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   465 AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGGTGGAGCTGCCGGCCG 524

Qy   203 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAAGCAGTGCCGTTGGTGTGCGAGGAGACGG 262
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAAGCAGTGCCGTTGGTGTGCGAGGAGACGG 262

```
XX
DR   WPI; 2002-557663/59.
XX
PT   Use of genes and their expression profiles associated with osteoblast
PT   differentiation for screening modulators bone formation, for diagnosing
PT   or treating e.g. osteoporosis, or as markers for the differentition
PT   process.
XX
PS   Claim 1; SEQ ID NO 9; 78pp + Sequence Listing; English.
XX
CC   The invention relates to genes and their expression profiles are used
CC   for: (a) screening modulators of precursor stem cell differentiation into
CC   osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC   deposition of bone tissue, abnormal rate of osteoblast formation or
CC   osteoporosis; or (c) treating or monitoring treatment of the conditions
CC   cited in (b), or monitoring the progression of bone tissue deposition.
CC   Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC   osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC   induced abnormalities in bone formation or bone loss, conditions that
CC   involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC   skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC   or fibrous dysplasia. The present sequence is that of an osteoblast
CC   differentiation associated cDNA marker of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Query Match             90.0%;  Score 772;  DB 6;  Length 1633;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches  822;  Conservative    0;  Mismatches   1;  Indels    0;  Gaps     0;

Qy    23  CCGGCCTTCGGCCGGCCACCCGACCCGTGGTCGGGCGCCACCCAGCGCGTGCGGGCGCTACCCGAGTCGCTCTGCCAGC   82
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   345  CCGGCCTTCGGCCGGCCACCCGACCCGTGGTCGGGCGCCACCCAGCGCGTGCGGGCGCTACCCGAGTCGCTCTGCCAGC  404

Qy    83  ACGGCGTGAGAAGCGCCAAGGCGAGGGGCGAGGAGGGTGGACGTGCGCCGCGGCGGACGCGGACGCAGCACC  142
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   405  ACGGCGTGAGAAGCGCCAAGGCGAGGGGCGAGGAGGGTGGACGTGCGCCGCGGCGGACGCGGACGCAGCACC  464

Qy   143  AGCTCTACGTGGCGGCTGGTGCAGCAAGCTGGGGCTGGAGCTGGGCTGCGGCCG  202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   465  AGCTCTACGTGGCGGCTGGTGCAGCAAGCTGGGGCTGGAGCTGGGCTGCGGCCG  524

Qy   203  ACGAGAGCCTTCCGGACCTGCCTTCGTCGTGGAGGACGTGGGCCGTGGTGTGCGGAGGACGG  262
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   525  ACGAGAGCCTTCCGGACCTGCCTTCGTCGTGGAGGACGTGGGCCGTGGTGTGCGGAGGACGG  584

Qy   263  CCCTCATCACCCGACCCGGCGGCGGCCGGAGGAAGGAGGTTGACATGATGAAAGAAG  322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   585  CCCTCATCACCCGACCCGGCGGCGGCCGGAGGAAGGAGGTTGACATGATGAAAGAAG  644

Qy   323  CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAAGATGAAAATGCAACTTTAGATG  382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   645  CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAAGATGAAAATGCAACTTTAGATG  704

Qy   383  GCGGAGATGTTTATTCACAGGCAGCAGAATTTTTGTGGGGCCTTTCCAAAAGGACAAATC  442
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   705  GCGGAGATGTTTATTCACAGGCAGCAGAATTTTTGTGGGGCCTTTCCAAAAGGACAAATC  764

Qy   443  AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATCAGTCTCCACAGTGCCAG  502
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   765  AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATCAGTCTCCACAGTGCCAG  824

Qy   503  TGGCAGATGGGTGCTGCATTTGAAGAGTTTCTGCAGCATGCTGGGCCTAACCTGATCGCAA  562
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   825  TGGCAGATGGGTGCTGCATTTGAAGAGTTTCTGCAGCATGCTGGGCCTAACCTGATCGCAA  884

Qy   563  TTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATCGCAACAGATGAGTGCCACC  622
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   885  TTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATCCAACAGATGAGTGCCACC  944

Qy   623  GCTACGACAAACTCACTGCGCTGATGACATAGCAGCAAACTGTATATATCCTAAATATCC  682
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   945  GCTACGACAAACTCACTGCGCTGATGACATAGCAGCAAACTGTATATATCCTAAATATCC 1004

Qy   683  CCAACAAAGGGCACGTCTTGCTCGTGCTGCTGCACCGGAACCCGGAAGAGTATCCAGAAAGTGCAAAGG  742
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1005  CCAACAAAGGGCACGTCTTGCTCGTGCTGCTGCACCGGAACCCGGAAGAGTATCCAGAAAGTGCAAAGG 1064

Qy   743  TTTATGAGAAACTGAAGGACCATATGTGATCCCGTGAGCATGTCTGAACTGGAAAAGG  802
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1065  TTTATGAGAAACTGAAGGACCATATGTGATCCCGTGAGCATGTCTGAACTGGAAAAGG 1124

Qy   803  TGGATGGGGTGCTCACCTGCTGCTCAGTTTAATTAACAAGAA  845
          ||||||||||||||||||||||||||||||||||||||||||
Db  1125  TGGATGGGGTGCTCACCTGCTGCTCAGTTTAATTAACAAGAA 1167
```

```
RESULT 4
ABQ88102
ID   ABQ88102 standard; cDNA; 1633 BP.
XX
AC   ABQ88102;
XX
DT   18-SEP-2002  (first entry)
XX
DE   Human osteoblast differentiation related cDNA SEQ ID NO 9.
XX
KW   Human; osteoblast; stem cell differentiation; bone tissue deposition;
KW   osteoporosis; osteopathic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200250301-A2.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-US048276.
XX
PR   18-DEC-2000; 2000US-0255882P.
PR   24-APR-2001; 2001US-0285691P.
XX
PA   (GENE-) GENE LOGIC INC.
PA   (PROC ) PROCTER & GAMBLE CO.
XX
PI   Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI   Mertz L;
```

US-09-889-753B-1.txt

RESULT 5
AB235939

ID	AB235939 standard; cDNA; 3932 BP.
AC	AB235939;
XX	
DT	10-FEB-2003 (first entry)
XX	
DE	Human secretory polynucleotide SPTM SEQ ID NO 103.
XX	
KW	Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
KW	asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
KW	Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
KW	multiple sclerosis; Parkinson's disease; cell proliferative disorder;
KW	anti-inflammatory; immunosuppressive; neuroprotective; nootropic;
KW	neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
KW	antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
KW	secretory polynucleotide; secretory protein; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200283876-A2.
XX	
PD	24-OCT-2002.
XX	
PF	27-MAR-2002; 2002WO-US009921
XX	
PR	29-MAR-2001; 2001US-0280067P.
PR	29-MAR-2001; 2001US-0280068P.
PR	16-MAY-2001; 2001US-0291280P.
PR	17-MAY-2001; 2001US-0291829P.
PR	17-MAY-2001; 2001US-0291849P.
PR	19-JUN-2001; 2001US-0299428P.
PR	20-JUN-2001; 2001US-0299776P.
PR	20-JUN-2001; 2001US-0300001P.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI	Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI	Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI	Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI	Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX	
DR	WPI; 2003-075543/07.
DR	P-PSDB; ABP75492.
XX	
PT	New human secretory proteins and polynucleotides, useful for diagnosing,
PT	treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
PT	neurological disorders (e.g. Alzheimer's), or cell proliferations or
PT	cancers.
XX	
PS	Claim 1; SEQ ID NO 103; 458pp + Sequence Listing; English.
XX	
CC	The invention relates to a secretory polynucleotide (designated sptm)
CC	comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
CC	naturally occurring polynucleotide sequence at least 90 % identical to
CC	the polynucleotide sequence, a polynucleotide complementary to them or an
CC	RNA equivalent of them. The polypeptide or polynucleotide are useful for
CC	treating, preventing or diagnosing a disease or condition associated with
CC	the expression of functional SPTM. These are particularly useful for
CC	diagnosing, treating or preventing autoimmune/inflammatory disorders
CC	(e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
CC	disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC	dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
CC	multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
CC	schizophrenia or amnesia), or cell proliferative disorders (e.g.
CC	psoriasis, polycythemia vera, or cancers including adenocarcinoma,
CC	leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
CC	breast, cervix or prostate). Note: The sequence data for this patent did
CC	not form part of the printed specification, but was obtained in electronic
CC	format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 3932 BP; 1089 A; 883 C; 775 G; 1185 T; 0 U; 0 Other;

Query Match
Best Local Similarity 90.0%; Score 772; DB 7; Length 3932;
Matches 822; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

```
Qy     23 CGGCCTTCGGCGGCCACCACGCCGTTGCGGCGCTACCCGAGTTCGCTCTGCCAGC  82
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     39 CGGCCTTCGGCGGCCACCACGCCGTGGCGCGGCGCTACCCGAGTCGCTCTGCCAGC  98

Qy     83 ACGCGCTGAGAGCGCCAAGGCGCGAGGAGTGGACGTCGCCGCGGAACGGCAGCACC 142
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99 ACGCGCTGAGAGCGCCAAGGCGCGAGGAGTGGACGTCGCCGCGGAACGGCAGCACC 158

Qy    143 AGTCTCACGTGGGCGTGCTGGGGCAGCAAGCTCGGGGTCCAGGTGGTGGAGCTCGCCGCG 202
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    159 AGTCTCACGTGGGCGTGCTGGGGCAGCAAGCTCGGGGTCCAGGTGGTGGAGCTCGCCGCG 218

Qy    203 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACGG 262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    219 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACGG 278

Qy    263 CCCTCATCACCCGACCCGGGGCCGGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAAGAAG 322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    279 CCCTCATCACCCGACCCGGGGCCGGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAAGAAG 338

Qy    323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATG 382
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    339 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATG 398

Qy    383 GCGGAGATGTTTATTCACACAGGCAGAGAATTTTTGTGGGCCTTTCAAAACGACAAATC 442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    399 GCGGAGATGTTTATTCACACAGGCAGAGAATTTTTGTGGGCCTTTCAAAACGACAAATC 458

Qy    443 AACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAG 502
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    459 AACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAG 518

Qy    503 TGGCAGATGGGTGCATTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAA 562
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    519 TGGCAGATGGGTGCATTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAA 578

Qy    563 TTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGCAGCAAACCACC 622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    579 TTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGCAGCAAACCACC 638

Qy    623 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 682
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    639 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 698

Qy    683 CCAACAAAGGGCACGTCTTGCTCACCGAACCCCGGAAGATCCAGAAAGTGCAAGG 742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    699 CCAACAAAGGGCACGTCTTGCTCACCGAACCCCGGAAGATCCAGAAAGTGCAAGG 758

Qy    743 TTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCTGAACTGGAAAGG 802
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Qy   623 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 682
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   945 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 1004

Qy   683 CCAACAAAGGGCACGTCTTGCTCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAGG 742
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1005 CCAACAAAGGGCACGTCTTGCTCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAGG 1064

Qy   743 TTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCTGAACTGGAAAAGG 802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1065 TTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCTGAACTGGAAAAGG 1124

Qy   803 TGGATGGGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA 845
          ||||||||||||||||||||||||||||||||||||||||
Db   1125 TGGATGGGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA 1167

```
Db    759 TTTATGAGAAACTGAAGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGG 818
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    803 TGGATGGGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA 845
          ||||||||||||||||||||||||||||||||||||||||
Db    819 TGGATGGGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA 861
```

RESULT 6
AAL41494
ID   AAL41494 standard; DNA; 4208 BP.
XX
AC   AAL41494;
XX
DT   21-NOV-2002   (first entry)
XX
DE   Drug metabolising enzyme encoding DNA - 7671089CB1.
XX
KW   Anti-HIV; antiarteriosclerotic; dermatological; cytostatic; thyromimetic;
KW   osteopathic; thrombolytic; ophthalmological; antilipaemic; hepatotropic;
KW   antidiarrhoeic; antiinflammatory; virucide; immunogenic; autoimmune;
KW   drug metabolising enzyme; DME; inflammatory; AIDS; atherosclerosis;
KW   contact dermatitis; cell-proliferative; cancer; cirrhosis; dwarfism;
KW   developmental; hypothyroidism; endocrine; osteoporosis; thrombosis;
KW   diabetes; glaucoma; keratitis; metabolic; hyperlipidaemia; diarrhoea;
KW   cystic fibrosis; gastrointestinal; gastroenteritis; liver; hepatitis; gene;
KW   Reye's syndrome; exogenous compound; gene therapy; enzyme; human; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   200..1225
FT   /tag= a
FT   /product= "Drug metabolising enzyme protein"
XX
PN   WO200266654-A2.
XX
PD   29-AUG-2002.
XX
PF   14-FEB-2002; 2002WO-US004918.
XX
PR   16-FEB-2001; 2001US-0269643P.
PR   23-FEB-2001; 2001US-0271332P.
PR   16-MAR-2001; 2001US-0276767P.
PR   06-APR-2001; 2001US-0282077P.
PR   19-APR-2001; 2001US-0285447P.
PR   27-APR-2001; 2001US-0287060P.
PR   03-MAY-2001; 2001US-0288543P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Astromoff A, Au-Young J, Baughn MR, Ding L, Duggan BM;
PI   Forsythe IJ, Gietzen KJ, Griffin JA, Lee EA, Lu Y, Richardson TW;
PI   Ring HZ, Sanjanwala MM, Swarnakar A, Walia NK, Warren BA, Xu Y;
PI   Yue H, Zebarjadian Y;
XX
DR   WPI; 2002-674949/72.
DR   P-PSDB; AAO22798.
XX
PT   New drug metabolizing enzymes (DME) useful for diagnosing, treating and
PT   preventing diseases or conditions associated with aberrant DME
PT   expression, e.g. cancer, AIDS, atherosclerosis, diabetes, glaucoma,
PT   hepatitis, osteoporosis.
XX
PS   Claim 5; Page 164-165; 166pp; English.
XX
CC   The invention relates to an isolated polypeptide comprising 12 fully
CC   defined sequences of 81-615 amino acids given in the specification; a
CC   naturally occurring amino acid sequence at least 90% identical to, having
CC   81-599 amino acids, at least 93% identical to a sequence of 529 amino
CC   acids, or at least 97% identical to a sequence of 615 amino acids, all
CC   given in the specification; or a biologically active or immunogenic
CC   fragment of the polypeptide. The polypeptides and polynucleotides are
CC   useful in diagnosing, treating and preventing diseases or conditions
CC   associated with the decreased expression or overexpression of a drug
CC   metabolising enzyme (DME), such as autoimmune/inflammatory (e.g. AIDS,
CC   atherosclerosis, contact dermatitis) cell-proliferative (e.g. cancer,
CC   cirrhosis), developmental (e.g. dwarfism, hypothyroidism), endocrine
CC   (e.g. osteoporosis, thrombosis, diabetes), eye (e.g. glaucoma,
CC   keratitis), metabolic (e.g. hyperlipidaemia, cystic fibrosis),
CC   gastrointestinal (e.g. gastroenteritis, diarrhoea), or liver (e.g.
CC   hepatitis, Reye's syndrome) disorders. These are also useful in assessing
CC   the effects of exogenous compounds on the expression of nucleic acid and
CC   amino acid sequences of DME. The DME or its fragments are useful in
CC   screening compounds for effectiveness as agonist or antagonist of the
CC   polypeptides, or in altering the expression of the target polynucleotide
CC   and compounds that specifically bind to or modulate the activity of the
CC   polypeptide. The microarray is useful in monitoring or measuring protein-
CC   protein interactions, drug-target interactions, and gene expression
CC   profiles. The polynucleotides of the invention can be used in gene
CC   therapy. This polynucleotide sequence represents the DNA encoding a drug
CC   metabolising enzyme of the invention
XX
SQ   Sequence 4208 BP; 1092 A; 1017 C; 883 G; 1216 T; 0 U; 0 Other;

Query Match   90.0%; Score 772; DB 6; Length 4208;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 822; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     23 CCGCCTTCGGCCGCGGCCACCCACGCCCGTGGTGCGGCGGCCTCACCCGAGTCGCTCTGCCAGC 82
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 CCGCCTTCGGCCGCGGCCACCCACGCCCGTGGTGCGGCGGCCTCACCCGAGTCGCTCTGCCAGC 449

Qy     83 ACGCGCTGAGAAGCGCCAAGGGCGGAGAGCCGGAGGAGGTTGACGTCGCCCGGCGGAGCAGCACC 142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450 ACGCGCTGAGAAGCGCCAAGGGCGGAGAGCCGGAGGAGGTTGACGTCGCCCGGCGGAGCAGCACC 509

Qy    143 AGCTTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCTGGTGAGCTGGTGGAGCTGCCGGCCG 202
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510 AGCTTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCTGGTGAGCTGGTGGAGCTGCCGGCCG 569

Qy    203 ACGAGAGCCTTCCGGACTGCGTCTTCGTCGGACGGTGCCGTGGTGTGCGAGGAGACGG 262
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    570 ACGAGAGCCTTCCGGACTGCGTCTTCGTCGGAGACGGTGCCGTGGTGTGCGAGGAGACGG 629

Qy    263 CCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAAGAAG 322
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    630 CCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAAGAAG 689

Qy    323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGAGAGATGAAAGATGAAATGCAACTTTAGATG 382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    690 CATTAGAAAAACTTCAGCTCAATATAGTAGAGAGAGATGAAAGATGAAATGCAACTTTAGATG 749

Qy    383 GCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC 442
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    750 GCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC 809

Qy    443 AACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGGACTCCACAGTGCCAG 502
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    810 AACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGGACTCCACAGTGCCAG 869

Qy    503 TGGCAGATGGGTTGCATTGAAGAGTTCTGCAGCATGGCTGGGCATGCGGCCTAACCTGATCGCAA 562
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    870 TGGCAGATGGGTTGCATTGAAGAGTTCTGCAGCATGGCTGGGCATGCGGCCTAACCTGATCGCAA 929

Qy    563 TTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACC 622
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    930 TTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACC 989

Qy    623 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 682
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    990 GCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 1049

Qy    683 CCAACAAAGGCACGCACGTCTTTGCTGCACCGACCCCGGAAGTGATCCGAAGTGCAAAGG 742
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 CCAACAAAGGCACGCACGTCTTTGCTGCACCGACCCCGGAAGTGATCCGAAGTGCAAAGG 1109
```

Qy   743 TTTATGAGAAACTGAAGGACCATATGTCGATCCCCCGTGAGCATGTCTGAACTGGAAAAGG 802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1110 TTTATGAGAAACTGAAGGACCATATGTCGATCCCCCGTGAGCATGTCTGAACTGGAAAAGG 1169

Qy   803 TGGATGGGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA 845
          ||||||||||||||||||||||||||||||||||||||||
Db 1170 TGGATGGGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA 1212

RESULT 7
ADE71463
ID   ADE71463   standard; DNA; 858 BP.
XX
AC   ADE71463;
XX
DT   29-JAN-2004   (first entry)
XX
DE   Human Met87 variant dimethylarginine dimethylaminohydrolase-1 DNA.
XX
KW   Human; variant dimethylarginine dimethylaminohydrolase-1; variant DDAH 1;
KW   chromosome 1; single nucleotide polymorphism; SNP; ADMA;
KW   asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
KW   coronary heart disease; cerebrovascular disease; hypertension; diabetes;
KW   susceptibility; genotyping; therapeutic targeting; antidiabetic;
KW   cardiant; cerebroprotective; hypotensive; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..858
FT   /*tag= a
FT   /product= "Met87 variant DDAH 1"
FT   replace(260,C)
FT   /*tag= b
FT   variation
FT   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   /note= "Results in Thr (C variant) or Met (T variant) at
FT   position 87 of the protein"
XX
PN   WO200309638-A1.
XX
PD   30-OCT-2003.
XX
PF   11-APR-2003; 2003WO-FI000274.
XX
PR   19-APR-2002; 2002US-0125456.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Valkonen V, Salonen JT, Pirskanen M, Tuomainen T, Laakso J;
PI   Laaksonen R;
XX
DR   WPI; 2003-854121/79.
DR   P-PSDB; ADE71464.
XX
PT   New nucleic acid encoding a variant dimethylarginine
PT   dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
PT   and its vascular complications, e.g. coronary or cerebrovascular disease
PT   or hypertension.
XX
PS   Claim 4; SEQ ID NO 1; 37pp; English.
XX
CC   The invention relates to a variant dimethylarginine
CC   dimethylaminohydrolase-1 (DDAH 1; ADE71464) and a nucleic acid encoding
CC   it (ADE71463). DDAHs regulate the metabolism of ADMA (asymmetrical
CC   dimethyl-arginine; a naturally occurring nitric oxide synthase inhibitor)
CC   and nitric oxide, which are relevant in disorders such as cardiovascular
CC   disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
CC   has been implicated in susceptibility to diseases such as familial
CC   combined hyperlipidaemia, premature coronary artery disease (CAD, non
CC   insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
CC   The variant DDAH 1 of the invention has Met at position 87, rather than
CC   Thr, which is caused by a C to T change in exon 1 at a position

CC   corresponding to base 260 in ADE71463 and ADE71465. Individuals hetero-
CC   or homozygous for the Thr87Met allele have an increased risk of
CC   developing cardiovascular diseases and diabetes. The invention also
CC   relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a
CC   susceptibility to a cardiovascular disease and diabetes in an individual
CC   by determining the DDAH 1 genotype; a method of treating diabetes or
CC   vascular complications of diabetes by enhancing nitric oxide
CC   availability, production or concentration; a method of targeting
CC   treatment of cardiovascular disease and diabetes in a hypertensive
CC   patient by determining the DDAH 1 genotype and treating them with a drug
CC   that affects nitric oxide availability, production or metabolism; and a
CC   transgenic animal which carries a human variant DDAH 1 nucleic acid
CC   sequence. The nucleic acid molecules and polypeptides are useful for
CC   treating cardiovascular disease (e.g., coronary heart disease,
CC   cerebrovascular disease, and hypertension), and diabetes and its vascular
CC   complications. The methods are useful for determining whether a patient
CC   will benefit from treatment with a drug which affects nitric oxide
CC   availability, production or metabolism; or an agent which elevates DDAH
CC   availability or concentration (such as DDAH agonist). The methods are
CC   also useful for determining whether a patient will be at risk of adverse
CC   effects if DDAH antagonists are administered. The present sequence
CC   represents a specifically claimed nucleic acid encoding the variant
CC   (Met87) DDAH 1.
XX
SQ   Sequence 858 BP; 218 A; 212 C; 258 G; 170 T; 0 U; 0 Other;

Query Match   84.0%;   Score 721;   DB 9;   Length 858;
Best Local Similarity   99.8%;   Pred. No. 0;
Matches 821;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

Qy   23 CGGCCTTCGGCCGGCCACCCACGCCGTGGTGCGGGGCGCTACCCGAGTTCGCTCTCGCCAGC 82
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   23 CGGCCTTCGGCCGGCCACCCACGCCGTGGTGCGGGGCGCTACCCGAGTTCGCTCTCGCCAGC 82

Qy   83 ACGGGCTGAGAAGCGCACAGGGCGAGGAGGTGGACGTCGCCGCCGCGGGCAGCGGCAGCACC 142
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83 ACGGGCTGAGAAGCGCACAGGGCGAGGAGGTGGACGTCGCCGCCGCGGGCAGCGGCAGCACC 142

Qy 143 AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGCCGGCGG 202
         ||||||||||||||||||||||||||||||||||||||||||||
Db 143 AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGCCGGCCG 202

Qy 203 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACGG 262
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 203 ACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGATGG 262

Qy 263 CCCTCATCACCCGACCCGGGCGCGCCGAGCCCGGAGGAAGGAGGTTGACATGATGAAAGAAG 322
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263 CCCTCATCACCCGACCCGGGCGCGCCGAGCCCGGAGGAAGGAGGTTGACATGATGAAAGAAG 322

Qy 323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAGAAATGCAACTTTAGATG 382
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAGAAATGCAACTTTAGATG 382

Qy 383 GCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC 442
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 383 GCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC 442

Qy 443 AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTCATGCAGTCTCCACAGTGCCAG 502
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 443 AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTCATGCAGTCTCCACAGTGCCAG 502

Qy 503 TGGCAGATGGGTTGCATTGAAGAGTTCTGCAGATGCTGGGGCTAACCTGATCGCAA 562
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 503 TGGCAGATGGGTTGCATTGAAGAGTTCTGCAGATGCTGGGGCTAACCTGATCGCAA 562

Qy 563 TTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCACC 622
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 563 TTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCACC 622

Qy 623 GCTACGACAAACTCACTGTGCCTGATGACATAGCAAACTGTATATATCCTAAATATCC 682
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Query Match   71.1%; Score 610; DB 3; Length 3651;
Best Local Similarity 100.0%; Pred. No. 5.1e-304;
Matches 610; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    215 CGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGGAGGAGACGGGCCCTCATCACCC 274
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 CGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGGAGGAGACGGGCCCTCATCACCC 61

Qy    275 GACCCGGGGCGCCGACCGGAGGAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAAC 334
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 GACCCGGGGCGCCGACCGGAGGAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAAAC 121

Qy    335 TTCAGCTCAATATAGTAGGATAGAGATGAAAATGAAAATGCAACTTTAGATGCGGAGATGTTT 394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 TTCAGCTCAATATAGTAGGATAGAGATGAAAATGAAAATGCAACTTTAGATGCGGAGATGTTT 181

Qy    395 TATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAAATCAACGAGGTGCTG 454
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 TATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAAATCAACAGGTGCTG 241

Qy    455 AAATCTTGGCTGATACTTTAAGGACTATGCAGTCCCACAGTGCCAGTGGCGCAGATGGGT 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 AAATCTTGGCTGATACTTTAAGGACTATGCAGTCCCACAGTGCCAGTGGCAGATGGGT 301

Qy    515 TGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTG 574
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 TGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTG 361

Qy    575 AATCTGCACAGAGGCCCTTAAGTACATCATGCAACAGCAGATGAGTGACCACCGCTACGACAAAC 634
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 AATCTGCACAGAGGCCCTTAAGATCATCATGCAACAGCAGATGAGTGACCACCGCTACGACAAAC 421

Qy    635 TCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCAACAAAGGGC 694
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 TCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCAACAAAGGGC 481

Qy    695 ACGTCTTGCTGCACCGACCCCGGAAGAGTATCCAGAAAGTTTATGAGAAAC 754
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 ACGTCTTGCTGCACCGACCCCGGAAGAGTATCCAGAAAGTTTATGAGAAAC 541

Qy    755 TGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGATGGGCTGC 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 TGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGATGGGCTGC 601

Qy    815 TCACCTGCTG 824
          ||||||||||
Db    602 TCACCTGCTG 611
```

RESULT 8
AAF16077
ID   AAF16077 standard; cDNA; 3651 BP.
XX
AC   AAF16077;
XX
DT   13-MAR-2001 (first entry)
XX
DE   Human prostate cancer antigen nucleotide sequence SEQ ID NO:512.
XX
KW   Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW   neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW   vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW   antibacterial; gene therapy; neural; immune; reproductive; renal;
KW   gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW   wound; infectious disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055174-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005988.
XX
PR   12-MAR-1999; 99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2000-587513/55.
DR   P-PSDB; AAB56874.
XX
PT   Prostate cancer associated gene sequences, referred to as prostate cancer
PT   antigens, useful for treatment, prevention, and diagnosis of disorders
PT   such as prostate cancer.
XX
PS   Claim 1; Page 976-977; 2338pp; English.
XX
CC   AAF15566 to AAF16505 encode the human prostate cancer associated
CC   proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC   The prostate cancer antigens can have neuroprotective, cytostatic,
CC   cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC   nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC   and can be used in gene therapy. The prostate cancer antigen
CC   polynucleotides may be used for detection of prostate cancer, chromosome
CC   identification as chromosome markers, and for numerous other diagnostic
CC   or research purposes. The prostate cancer antigens may be used to treat
CC   disorders such as neural, immune, muscular, reproductive, renal and proliferative
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC   AAB57303 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 3651 BP; 1016 A; 801 C; 684 G; 1141 T; 0 U; 9 Other;

```
Db    623 GCTACGACAAACTCACTGCTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    683 CCAACAAAGGCACGCTTGCTGCTGCTGCCGACCCCGGAAGAGTATCCAGAAAGTGCAAAGG 742
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 CCAACAAAGGGCACGCTCTGCTGCCGACCCCGGAAGAGTATCCAGAAAGTGCAAAGG 742

Qy    743 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCGAACTGGAAAAGG 802
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGG 802

Qy    803 TGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAACAAGAA 845
          |||||||||||||||||||||||||||||||||||||||||||
Db    803 TGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAACAAGAA 845
```

RESULT 9
ADE53981
ID   ADE53981 standard; cDNA; 3778 BP.
XX
AC   ADE53981;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human prostate cancer cDNA #328.
XX
KW   Human; prostate cancer; ss; cDNA combination; differential expression;
KW   gene.
XX
OS   Homo sapiens.
XX
PN   US2003190640-A1.
XX
PD   09-OCT-2003.
XX
PF   29-MAY-2002; 2002US-00252157.
XX
PR   31-MAY-2001; 2001US-0295048P.
XX
PA   (FARI/) FARIS M.

PA   (PEAR/) PEARSON C I.
XX
PI   Faris M, Pearson CI;
XX
DR   WPI; 2003-831619/77.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   prostate cancer, useful for diagnosing, treating or monitoring the
PT   progression of treatment of prostate cancer.
XX
PS   Claim 1; SEQ ID NO 328; 42pp; English.
XX
CC   The invention relates to a combination comprising a number of cDNAs
CC   expressed in prostate cancer. The invention also relates to a method for
CC   detecting differential expression of one or more cDNAs in a sample
CC   containing nucleic acids by hybridising a substrate with the nucleic
CC   acids, thus forming one or more hybridisation complexes, detecting
CC   hybridisation complex formation and comparing the complexes formed with
CC   standard complexes, where differences between the standard and the sample
CC   complex formation indicate differential expression of cDNAs in the
CC   sample. The differential expression is diagnostic of prostate cancer. The
CC   invention also relates to proteins and antibodies related to the cDNAs.
CC   The combination is useful for diagnosing, treating or monitoring the
CC   progression of treatment of prostate cancer. The antibodies are useful
CC   for detecting prostate cancer. This sequence represents a human prostate
CC   cancer cDNA of the invention.
XX
SQ   Sequence 3778 BP; 1064 A; 821 C; 719 G; 1150 T; 0 U; 24 Other;

Query Match   68.2%;   Score 585;   DB 9;   Length 3778;
Best Local Similarity   99.9%;   Pred. No. 4.1e-291;
Matches 705;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   1;

```
Qy    141  CCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCCTGCAGGTGGTGGGACTGCCGGC  200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  CCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCCTGCAGGTGGTGGGACTGCCGGC   60

Qy    201  CGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGAC  260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGAC  120

Qy    261  GGGCCTCATCACCCGACCCGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAAGA   320
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GGGCCTCATCACCCGACCCGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAAGA   180

Qy    321  AGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAATGCAACTTTAGA       380
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAATGCAACTTTAGA       240

Qy    381  TGGCGGAGATGTTTATTCACAGGCAGGAGAATTTTGTGGGCCTTTCCAAAAGGACAAA    440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  TGGGGGAGATGTTTATTCACAGGCAGGAGAATTTTGTGGGCCTTTCCAAAAGGACAAA    300

Qy    441  T-CAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGC   499
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  TCCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGC   360

Qy    500  CAGTGGCAGATGGGTTGCATTGAAGAGTTCTGCAGCATGGTGGGCTAACCTGATCG     559
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  CAGTGGCAGATGGGTTGCATTGAAGAGTTCTGCAGCATGGTGGGCTAACCTGATCG     420

Qy    560  CAATTGGGTTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACC  619
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  CAATTGGGTTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACC  480

Qy    620  ACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATCTAAATA    679
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  ACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATCTAAATA    540

Qy    680  TCCCCAACAAAGGGCACGTCTTGCTGCACCGGAACCCCGAAGAGTATCCGAAAGTGCAA   739
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  TCCCCAACAAAGGGCACGTCTTGCTGCACCGGAACCCCGAAGAGTATCCAGAAAGTGCAA  600
```

Qy   740   AGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGAGCGATGTCTGAACTCGAAA   799
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601   AGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGAGCGATGTCTGAACTCGAAA   660

Qy   800   AGGTGGATGGGCTGGCTCACCTGCTCAGTTTTAATTAACAAGAA   845
            ||||||||||||||||||||||||||||||||||||||||||||
Db   661   AGGTGGATGGGCTGGCTCACCTGCTCAGTTTTAATTAACAAGAA   706

RESULT 10
ABQ88101/c
ID   ABQ88101 standard; cDNA; 83946 BP.
AC   ABQ88101;
XX
DT   18-SEP-2002   (first entry)
XX
DE   Human osteoblast differentiation related cDNA SEQ ID NO 8.
XX
KW   Human; osteoblast; stem cell differentiation; bone tissue deposition;
KW   osteoporosis; osteopathic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200250301-A2.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-US048276.
XX
PR   18-DEC-2000; 2000US-0255882P.
PR   24-APR-2001; 2001US-0285691P.
XX
PA   (GENE-) GENE LOGIC INC.
PA   (PROC) PROCTER & GAMBLE CO.
XX
PI   Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI   Mertz L;
XX
DR   WPI; 2002-557663/59.
XX
PT   Use of genes and their expression profiles associated with osteoblast
PT   differentiation for screening modulators bone formation, for diagnosing
PT   or treating e.g. osteoporosis, or as markers for the differentiation
PT   process.
XX
PS   Claim 1; SEQ ID NO 8; 78pp + Sequence Listing; English.
XX
CC   The invention relates to genes and their expression profiles are used
CC   for: (a) screening modulators of precursor stem cell differentiation into
CC   osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC   deposition of bone tissue, abnormal rate of osteoblast formation or
CC   osteoporosis; or (c) treating or monitoring treatment of the conditions
CC   cited in (b), or monitoring the progression of bone tissue deposition.
CC   Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC   osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC   induced abnormalities in bone formation or bone loss, conditions that
CC   involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC   skeletal disease linked to breast cancer, mastocyosis, Fanconi syndrome
CC   or fibrous dysplasia. The present sequence is that of an osteoblast
CC   differentiation associated cDNA marker of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp:wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 83946 BP; 26139 A; 16620 C; 16296 G; 24891 T; 0 U; 0 Other;

Query Match   17.2%;   Score 148;   DB 6;   Length 83946;
Best Local Similarity   100.0%;   Pred. No. 1.3e-65;
Matches 148;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   596   AGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAG   655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   18612 AGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAG 18553
Qy     656 CAGCAAACTGTATATATCTAAATATTCCCAACAAAGGGCACGTCTTGCTGCCACCGAACCC 715
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   18552 CAGCAAACTGTATATATCTAAATATTCCCAACAAAGGGCACGTCTTGCTGCACCGAACCC 18493

Qy     716 CGGAAGAGTATCCAGAAAGTGCAAAGGT 743
             ||||||||||||||||||||||||||||
Db   18492 CGGAAGAGTATCCAGAAAGTGCAAAGGT 18465


RESULT 11
ABA14642
ID   ABA14642 standard; DNA; 6373 BP.
XX
AC   ABA14642;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Human nervous system related polynucleotide SEQ ID NO 6973.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001334.
XX
PR   31-JAN-2000;  2000US-0179065P.
PR   04-FEB-2000;  2000US-0180628P.
PR   24-FEB-2000;  2000US-0184664P.
PR   02-MAR-2000;  2000US-0186350P.
PR   16-MAR-2000;  2000US-0189874P.
PR   17-MAR-2000;  2000US-0190076P.
PR   18-APR-2000;  2000US-0198123P.
PR   19-MAY-2000;  2000US-0205515P.
PR   07-JUN-2000;  2000US-0209467P.
PR   28-JUN-2000;  2000US-0214886P.
PR   30-JUN-2000;  2000US-0215135P.
PR   07-JUL-2000;  2000US-0216647P.
PR   07-JUL-2000;  2000US-0216880P.
PR   11-JUL-2000;  2000US-0217487P.
PR   11-JUL-2000;  2000US-0217496P.
PR   14-JUL-2000;  2000US-0218290P.
PR   26-JUL-2000;  2000US-0220963P.
PR   26-JUL-2000;  2000US-0220964P.
PR   14-AUG-2000;  2000US-0224518P.
PR   14-AUG-2000;  2000US-0224519P.
PR   14-AUG-2000;  2000US-0225213P.
PR   14-AUG-2000;  2000US-0225214P.
PR   14-AUG-2000;  2000US-0225266P.
PR   14-AUG-2000;  2000US-0225267P.
PR   14-AUG-2000;  2000US-0225268P.
PR   14-AUG-2000;  2000US-0225270P.
PR   14-AUG-2000;  2000US-0225447P.
PR   14-AUG-2000;  2000US-0225757P.
PR   14-AUG-2000;  2000US-0225758P.
PR   14-AUG-2000;  2000US-0225759P.
PR   18-AUG-2000;  2000US-0226279P.
PR   22-AUG-2000;  2000US-0226681P.
PR   22-AUG-2000;  2000US-0226868P.
PR   22-AUG-2000;  2000US-0227182P.
PR   23-AUG-2000;  2000US-0227009P.
PR   30-AUG-2000;  2000US-0228924P.
PR   01-SEP-2000;  2000US-0229287P.
PR   01-SEP-2000;  2000US-0229343P.
PR   01-SEP-2000;  2000US-0229344P.
PR   01-SEP-2000;  2000US-0229345P.
PR   05-SEP-2000;  2000US-0229509P.
PR   05-SEP-2000;  2000US-0229513P.
PR   06-SEP-2000;  2000US-0230437P.
PR   06-SEP-2000;  2000US-0230438P.
PR   08-SEP-2000;  2000US-0231242P.
PR   08-SEP-2000;  2000US-0231243P.
PR   08-SEP-2000;  2000US-0231244P.
PR   08-SEP-2000;  2000US-0231413P.
PR   08-SEP-2000;  2000US-0231414P.
PR   08-SEP-2000;  2000US-0232080P.
PR   08-SEP-2000;  2000US-0232081P.
PR   12-SEP-2000;  2000US-0231968P.
PR   14-SEP-2000;  2000US-0232397P.
PR   14-SEP-2000;  2000US-0232398P.
PR   14-SEP-2000;  2000US-0232399P.
PR   14-SEP-2000;  2000US-0232400P.
PR   14-SEP-2000;  2000US-0232401P.
PR   14-SEP-2000;  2000US-0233063P.
PR   14-SEP-2000;  2000US-0233064P.
PR   14-SEP-2000;  2000US-0233065P.
PR   21-SEP-2000;  2000US-0234223P.
PR   21-SEP-2000;  2000US-0234274P.
PR   25-SEP-2000;  2000US-0234997P.
PR   25-SEP-2000;  2000US-0234998P.
PR   26-SEP-2000;  2000US-0235484P.
PR   27-SEP-2000;  2000US-0235834P.
PR   29-SEP-2000;  2000US-0358336P.
PR   29-SEP-2000;  2000US-0363327P.
PR   29-SEP-2000;  2000US-0363367P.
PR   29-SEP-2000;  2000US-0363368P.
PR   29-SEP-2000;  2000US-0363369P.
PR   29-SEP-2000;  2000US-0363370P.
PR   02-OCT-2000;  2000US-0368802P.
PR   02-OCT-2000;  2000US-0237037P.
PR   02-OCT-2000;  2000US-0237038P.
PR   02-OCT-2000;  2000US-0237039P.
PR   02-OCT-2000;  2000US-0237040P.
PR   13-OCT-2000;  2000US-0239935P.
PR   20-OCT-2000;  2000US-0239937P.
PR   20-OCT-2000;  2000US-0240960P.
PR   20-OCT-2000;  2000US-0241785P.
PR   20-OCT-2000;  2000US-0241786P.
PR   20-OCT-2000;  2000US-0241787P.
PR   20-OCT-2000;  2000US-0241808P.
PR   20-OCT-2000;  2000US-0241809P.
PR   20-OCT-2000;  2000US-0241826P.
PR   20-OCT-2000;  2000US-0242221P.
PR   01-NOV-2000;  2000US-0244617P.
PR   08-NOV-2000;  2000US-0246474P.
PR   08-NOV-2000;  2000US-0246475P.
PR   08-NOV-2000;  2000US-0246476P.
PR   08-NOV-2000;  2000US-0246477P.
PR   08-NOV-2000;  2000US-0246478P.
PR   08-NOV-2000;  2000US-0246523P.
PR   08-NOV-2000;  2000US-0246524P.
PR   08-NOV-2000;  2000US-0246525P.
PR   08-NOV-2000;  2000US-0246526P.
PR   08-NOV-2000;  2000US-0246527P.
PR   08-NOV-2000;  2000US-0246528P.
PR   08-NOV-2000;  2000US-0246532P.
PR   08-NOV-2000;  2000US-0246609P.
PR   08-NOV-2000;  2000US-0246610P.
PR   08-NOV-2000;  2000US-0246611P.
PR   08-NOV-2000;  2000US-0246613P.
PR   17-NOV-2000;  2000US-0249207P.
PR   17-NOV-2000;  2000US-0249208P.
PR   17-NOV-2000;  2000US-0249209P.
PR   17-NOV-2000;  2000US-0249210P.
PR   17-NOV-2000;  2000US-0249211P.
```

```
RESULT 12
ABA14641
ID   ABA14641 standard; DNA; 6381 BP.
XX
AC   ABA14641;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Human nervous system related polynucleotide SEQ ID NO 6972.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001334.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0203091P.
PR   01-DEC-2000; 2000US-0251160P.
PR   01-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   06-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-541565/60.
XX
PT   Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating nervous system cancers
PT   and metastases.
XX
PS   Disclosure; SEQ ID NO 6973; 1701pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC   (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 6373 BP; 1810 A; 1396 C; 1180 G; 1987 T; 0 U; 0 Other;

Query Match             13.3%;  Score 114;  DB 5;  Length 6373;
Best Local Similarity   100.0%;  Pred. No. 4.3e-48;
Matches  114;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    630 CAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCCAACAA 689
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCCAACAA 60

Qy    690 AGGGCACGTCTTCTGCCACCGACCCCCGGAAGAGTATCCAGAAAGTGCAAAGGT 743
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGGGCACGTCTTCTGCCACCGACCCCCGGAAGAGTATCCAGAAAGTGCAAAGGT 114
```

PR   08-SEP-2000;   2000US-0231413P.
PR   08-SEP-2000;   2000US-0231414P.
PR   08-SEP-2000;   2000US-0232080P.
PR   08-SEP-2000;   2000US-0232081P.
PR   12-SEP-2000;   2000US-0231968P.
PR   14-SEP-2000;   2000US-0233397P.
PR   14-SEP-2000;   2000US-0233398P.
PR   14-SEP-2000;   2000US-0233399P.
PR   14-SEP-2000;   2000US-0232400P.
PR   14-SEP-2000;   2000US-0232401P.
PR   14-SEP-2000;   2000US-0233063P.
PR   14-SEP-2000;   2000US-0233064P.
PR   14-SEP-2000;   2000US-0233065P.
PR   21-SEP-2000;   2000US-0234223P.
PR   21-SEP-2000;   2000US-0234274P.
PR   21-SEP-2000;   2000US-0234997P.
PR   25-SEP-2000;   2000US-0234998P.
PR   26-SEP-2000;   2000US-0235484P.
PR   27-SEP-2000;   2000US-0235834P.
PR   27-SEP-2000;   2000US-0235836P.
PR   29-SEP-2000;   2000US-0236327P.
PR   29-SEP-2000;   2000US-0236367P.
PR   29-SEP-2000;   2000US-0236368P.
PR   29-SEP-2000;   2000US-0236369P.
PR   29-SEP-2000;   2000US-0236370P.
PR   02-OCT-2000;   2000US-0236802P.
PR   02-OCT-2000;   2000US-0237037P.
PR   02-OCT-2000;   2000US-0237038P.
PR   02-OCT-2000;   2000US-0237039P.
PR   02-OCT-2000;   2000US-0237040P.
PR   13-OCT-2000;   2000US-0239935P.
PR   13-OCT-2000;   2000US-0239937P.
PR   20-OCT-2000;   2000US-0240960P.
PR   20-OCT-2000;   2000US-0241785P.
PR   20-OCT-2000;   2000US-0241786P.
PR   20-OCT-2000;   2000US-0241787P.
PR   20-OCT-2000;   2000US-0241808P.
PR   20-OCT-2000;   2000US-0241809P.
PR   20-OCT-2000;   2000US-0241826P.
PR   20-OCT-2000;   2000US-0242221P.
PR   01-NOV-2000;   2000US-0244617P.
PR   08-NOV-2000;   2000US-0246474P.
PR   08-NOV-2000;   2000US-0246475P.
PR   08-NOV-2000;   2000US-0246476P.
PR   08-NOV-2000;   2000US-0246478P.
PR   08-NOV-2000;   2000US-0246478P.
PR   08-NOV-2000;   2000US-0246523P.
PR   08-NOV-2000;   2000US-0246524P.
PR   08-NOV-2000;   2000US-0246525P.
PR   08-NOV-2000;   2000US-0246526P.
PR   08-NOV-2000;   2000US-0246527P.
PR   08-NOV-2000;   2000US-0246528P.
PR   08-NOV-2000;   2000US-0246532P.
PR   08-NOV-2000;   2000US-0246609P.
PR   08-NOV-2000;   2000US-0246610P.
PR   08-NOV-2000;   2000US-0246611P.
PR   08-NOV-2000;   2000US-0246613P.
PR   17-NOV-2000;   2000US-0249207P.
PR   17-NOV-2000;   2000US-0249208P.
PR   17-NOV-2000;   2000US-0249209P.
PR   17-NOV-2000;   2000US-0249210P.
PR   17-NOV-2000;   2000US-0249211P.
PR   17-NOV-2000;   2000US-0249212P.
PR   17-NOV-2000;   2000US-0249213P.
PR   17-NOV-2000;   2000US-0249214P.
PR   17-NOV-2000;   2000US-0249215P.
PR   17-NOV-2000;   2000US-0249216P.
PR   17-NOV-2000;   2000US-0249217P.
PR   17-NOV-2000;   2000US-0249218P.
PR   17-NOV-2000;   2000US-0249244P.
PR   17-NOV-2000;   2000US-0249245P.
PR   17-NOV-2000;   2000US-0249264P.
PR   17-NOV-2000;   2000US-0249265P.

PR   17-NOV-2000;   2000US-0249297P.
PR   17-NOV-2000;   2000US-0249299P.
PR   17-NOV-2000;   2000US-0249300P.
PR   01-DEC-2000;   2000US-0250391P.
PR   01-DEC-2000;   2000US-0251160P.
PR   05-DEC-2000;   2000US-0251030P.
PR   05-DEC-2000;   2000US-0251988P.
PR   05-DEC-2000;   2000US-0256719P.
PR   06-DEC-2000;   2000US-0251479P.
PR   08-DEC-2000;   2000US-0251856P.
PR   08-DEC-2000;   2000US-0251868P.
PR   08-DEC-2000;   2000US-0251869P.
PR   08-DEC-2000;   2000US-0251989P.
PR   08-DEC-2000;   2000US-0251990P.
PR   11-DEC-2000;   2000US-0254097P.
PR   05-JAN-2001;   2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,   Barash SC,   Ruben SM;
XX
DR   WPI; 2001-541565/60.
XX
PT   Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating nervous system cancers
PT   and metastases.
XX
PS   Disclosure; SEQ ID NO 6972; 1701pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC   (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 6381 BP; 1809 A; 1396 C; 1182 G; 1994 T; 0 U; 0 Other;

Query Match   13.3%; Score 114; DB 5; Length 6381;
Best Local Similarity   100.0%; Pred. No. 4.3e-48;
Matches 114;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   630 CAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAA 689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAA   60

Qy   690 AGGGCACGTCTTGCTGCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAAGGT 743
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGGGCACGTCTTGCTGCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAAGGT 114

RESULT 13
AAC10265
ID   AAC10265 standard; cDNA; 420 BP.
XX
AC   AAC10265;
XX
DT   06-OCT-2000 (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 14340.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PP   21-FEB-2000; 2000BP-00200610.
XX
PR   26-FEB-1999; 99US-0122487P.
XX
PA   (GBST) GENSET.
XX
PI   Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 14340; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 420 BP; 131 A; 95 C; 94 G; 97 T; 0 U; 3 Other;

Query Match   9.9%; Score 85; DB 3; Length 420;
Best Local Similarity   100.0%; Pred. No. 3.7e-33;
Matches   85; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   302 AGGTTGACATGATGAAGAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAG 361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   336 AGGTTGACATGATGAAGAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAG 395

Qy   362 ATGAAAATGCAACTTTAGATGGCGG 386
           |||||||||||||||||||||||||
Db   396 ATGAAAATGCAACTTTAGATGGCGG 420

RESULT 14
ABN36282
ID   ABN36282 standard; DNA; 60 BP.
XX
AC   ABN36282;
XX
DT   15-JUL-2002 (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:9030.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO20210449-A2.
XX
PD   07-FEB-2002.
XX

PF   20-JUL-2001; 2001WO-IB001903.
XX
PR   28-JUL-2000; 2000US-0221607P.
PR   02-MAY-2001; 2001US-0287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of a
PT   genome, useful for detecting tissue-, pathology-, and developmental-
PT   specific genes.
XX
PS   Example 1; SEQ ID NO 9030; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC   )transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises several
CC   oligonucleotides, each capable of hybridising selectively to a set of
CC   messenger RNAs transcribed from a given transcription unit of the genome,
CC   which encodes one or more messenger RNA splice variants. The
CC   oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a particular
CC   biological or pathological state, and so allowing the detection of tissue
CC   - and pathology-specific genes such as those genes only expressed in
CC   specific tissue under a specific pathological condition; to detect
CC   developmental specific genes; and to detect RNA transcripts and splice
CC   variants of a transcriptome of a patient suffering from a particular
CC   disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC   rats, humans and mice, which are used in the exemplification of the
CC   present invention. N.B. The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 60 BP; 18 A; 11 C; 19 G; 12 T; 0 U; 0 Other;

Query Match   7.0%; Score 60; DB 6; Length 60;
Best Local Similarity   100.0%; Pred. No. 2.8e-20;
Matches   60; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   751 AAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCGAACTGGAAAAGGTGGATGGG 810
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCGAACTGGAAAAGGTGGATGGG 60

RESULT 15
ADB53014
ID   ADB53014 standard; DNA; 3008 BP.
XX
AC   ADB53014;
XX
DT   04-DEC-2003 (first entry)
XX
DE   Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3556.
XX
KW   toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW   toxicity marker; toxicity progression; drug screening;
KW   primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO2003065993-A2.
XX
PD   14-AUG-2003.
XX

PF   04-FEB-2003; 2003WO-US003482.
XX
PR   04-FEB-2002; 2002US-0353171P.
PR   13-MAR-2002; 2002US-0363534P.
PR   08-APR-2002; 2002US-0370248P.
PR   10-APR-2002; 2002US-0371134P.
PR   10-APR-2002; 2002US-0371135P.
PR   10-APR-2002; 2002US-0371150P.
PR   11-APR-2002; 2002US-0371413P.
PR   19-APR-2002; 2002US-0373601P.
PR   19-APR-2002; 2002US-0373602P.
PR   22-APR-2002; 2002US-0374139P.
PR   08-MAY-2002; 2002US-0378370P.
PR   09-MAY-2002; 2002US-0378652P.
PR   09-MAY-2002; 2002US-0378653P.
PR   09-MAY-2002; 2002US-0378665P.
PR   09-JUL-2002; 2002US-0394230P.
PR   09-JUL-2002; 2002US-0394253P.
PR   04-SEP-2002; 2002US-0407688P.
PR   28-JAN-2003; 2003US-0442900P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,   Porter M,   Johnson K,   Higgs B,   Castle A,   Orr M;
PI   Elashoff M;
XX
DR   WPI; 2003-731472/69.
XX
PT   Determining if a compound induces a toxic effect on a tissue or cell, for
PT   identifying hepatotoxic compounds, comprises comparing a gene expression
PT   profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT   mean values.
XX
PS   Claim 44; SEQ ID NO 3556; 874pp; English.
XX
CC   The present invention describes a method for determining whether a
CC   compound induces a toxic effect on a tissue or cell. The method comprises
CC   preparing a gene expression profile of a tissue or cell sample exposed to
CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;

Query Match   6.3%; Score 54; DB 9; Length 3008;
Best Local Similarity   100.0%; Pred. No. 3.9e-17;
Matches   54; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   136 CAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGGTGGTG 189
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 CAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGGTGGTG 620

Search completed: June 7, 2004, 21:31:44
Job time : 407 secs

```
                Copyright   GenCore version 5.1.6
                Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   June  7, 2004, 20:34:35 ; Search time 2569 Seconds
                                    (without alignments)
                                    9973.434 Million cell updates/sec

Title:       US-09-889-733B-1
Perfect score: 858
Sequence:   1 atggccggcctcggccaccc........acaagaaggtagactcctga 858

Scoring table: OLIGO_NUC
               Gapop 60.0 , Gapext 60.0

Searched:    2751289 seqs, 1493109276 residues

Word size :  0

Total number of hits satisfying chosen parameters:       55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :    EST:*
         1:  em_estba:*
         2:  em_esthum:*
         3:  em_estin:*
         4:  em_estmu:*
         5:  em_estov:*
         6:  em_estpl:*
         7:  em_estro:*
         8:  em_htc:*
         9:  gb_est1:*
        10:  gb_est2:*
        11:  gb_htc:*
        12:  gb_est3:*
        13:  gb_est4:*
        14:  gb_est5:*
        15:  em_estfun:*
        16:  em_estom:*
        17:  em_gss_hum:*
        18:  em_gss_inv:*
        19:  em_gss_pln:*
        20:  em_gss_vrt:*
        21:  em_gss_fun:*
        22:  em_gss_mam:*
        23:  em_gss_mus:*
        24:  em_gss_pro:*
        25:  em_gss_rod:*
        26:  em_gss_phg:*
        27:  em_gss_vrl:*
        28:  gb_gss1:*
        29:  gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                    SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	516	60.1	522	9	AI751463	AI751463 cn10b04.x
2	457	53.3	889	12	BI763990	BI763990 603049879
3	442	51.5	455	29	AY406329	AY406329 Homo sapi
4	350	40.8	356	12	BG977444	BG977444 RC5-CI016

```
 5   340  39.6  455  29      AY406330   AY406330 Pan trogl
 6   323  37.6  624  10      AW514143   AW514143 hd24a02.x
 7   269  31.4  535   9      AU279491   AU279491 AU279491
 8   263  30.7  308  10      BF818963   BF818963 RC5-CI016
 9   216  25.2  406  10      AI017468   AI017468 ou23c07.x
10   159  18.5  276  10      BF818967   BF818967 RC5-CI016
11   138  16.1 1038  12      BG164988   BG164988 602343844
12   115  13.4  312   9      AA376335   AA376335 EST89005
13   106  12.4  700  28      AQ742098   AQ742098 HS_5566_B
14    59   6.9  167  10      BF933706   BF933706 CM2-NT017
15    59   6.9  357  13      BY233360   BY233360 BY233360
16    59   6.9  361  10      BB843865   BB843865 BB843865
17    59   6.9  387  13      BY292033   BY292033 BY292033
18    59   6.9  442  13      BY266349   BY266349 BY266349
19    59   6.9  498   9      AI930353   AI930353 u163f11.y
20    59   6.9  507   9      AI225556   AI225556 uj05e03.y
21    59   6.9  539   9      AW012296   AW012296 um07b05.y
22    59   6.9  550   9      AA986603   AA986603 uc81h06.y
23    59   6.9  555   9      AI050531   AI050531 uc83d09.y
24    59   6.9  638  13      BY713422   BY713422 BY713422
25    59   6.9  685  14      CB247154   CB247154 UI-M-FI0-
26    59   6.9  702  13      BY731938   BY731938 BY731938
27    59   6.9  731  10      BB611533   BB611533 BB611533
28    59   6.9  759  12      BI736750   BI736750 60336O448
29    59   6.9  777  14      CD349526   CD349526 UI-M-FY0-
30    59   6.9  806   9      AU051351   AU051351 AU051351
31    59   6.9  844   9      AU051035   AU051035 AU051035
32    59   6.9  862  13      BQ899468   BQ899468 AGENCOURT
33    59   6.9  923  12      BI332932   BI332932 602982027
34    59   6.9 1001  13      BY709864   BY709864 BY709864
35    59   6.9 1109  11      AK010430   AK010430 Mus muscu
36    55   6.4  224   9      AA376525   AA376525 EST88957
37    54   6.3  450  14      CB787749   CB787749 AMGNNUC:N
38    54   6.3  493  14      CB725087   CB725087 AMGNNUC:M
39    54   6.3  524  14      CB720352   CB720352 AMGNNUC:N
40    48   5.6  547   9      AI661738   AI661738 va31a06.x
41    47   5.5  874  12      BI560188   BI560188 603253311
42    47   5.5  338  13      BY331431   BY331431 BY331431
43    47   5.5  343   9      AA403666   AA403666 va31a06.r
44    47   5.5  346   9      AA499371   AA499371 vi89b12.r
45    47   5.5  388  13      BX490218   BX490218 DKFZp686H

                                              ALIGNMENTS

RESULT 1
AI751463
LOCUS       AI751463              522 bp    mRNA    linear    EST 20-JUN-2002
DEFINITION  cn10b04.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
            clone NHTBC_cn10b04 random, mRNA sequence.
ACCESSION   AI751463
VERSION     AI751463.1  GI:5129727
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
            Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
            Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
  TITLE     SGAP:The Skeletal Genome Anatomy Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Libin Jia
            Medical Genetics Branch
            National Human Genome Research Institute
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libing@helix.nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
```

TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11474 row: n column: 19
            High quality sequence start: 3
            High quality sequence stop: 847

FEATURES   Location/Qualifiers
     source   1..889
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5189970"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_116"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female. Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon cloning).   Average insert size 1.4 kb,
                     insert size range 1-3 kb.   Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber (Invitrogen).   Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."

ORIGIN

Query Match   53.3%;   Score 457;   DB 12;   Length 889;
Best Local Similarity 99.8%;   Pred. No. 1.2e-226;
Matches 507; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   23 CCGGCTTCGGCGGGCCACCGCCGTGGTGCGGGGCGCGTACCCGAGTCGCTCTGCCAGC 82
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 CCGCCTTCGGCGGGCCACCGCCGTGGTGCGGGGCGCGTACCCGAGTCGCTCGGCCAGC 183

Qy   83 ACGCGCTGAGAAGCCAAGGCCGGCCAAGGCGGAGGAGGTCGCCGGACGTCGGCAGCACC 142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 ACGCGCTGAGAAGCCCAAGGCGCAAGGCGGAGGAGGTGGCACGTCGCCGGCGGAGCAGCACC 243

Qy   143 AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGGTGGAGCTGCGGAGGACGGTCCG 202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 AGTCTCACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTGGTGGAGCTGCGGAGGACGGTCCG 303

Qy   203 ACGAGAGCCCTTCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGTGCGGAGGGACGG 262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304 ACGAGAGCCCTTCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGTGCGGAGGGACGG 363

Qy   263 CCCTCATCACCCGACCCGGGCGCCGAGGCAGGAGGTTGACATGATGAAAGAAG 322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 CCCTCATCACCCGACCCGGGCGCCGAGGCAGGAGGTTGACATGATGAAAGAAG 423

Qy   323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAGATGAAAATGCAACTTTAGATG 382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAGATGAAAATGCAACTTTAGATG 483

Qy   383 GCGGACGATGTTTATTCACAGGCAGCAGAATTTTTGTGGGGCCTTTCCAAAAGGACAAATC 442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484 GCGGACGATGTTTATTCACAGGCAGCAGAATTTTTGTGGGGCCTTTCCAAAAGGACAAATC 543

Qy   443 AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAG 502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544 AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAG 603

Qy   503 TGGCAGATGGGTTGCATTGAAGAGTTT 530
          |||||||||||||||||||||||||

FEATURES   Location/Qualifiers
     source   1..522
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NHTBC_cn10b04"
                     /sex="Female"
                     /tissue_type="Bone"
                     /cell_type="Trabecular Bone Cells"
                     /lab_host="SURE"
                     /clone_lib="Normal Human Trabecular Bone Cells"
                     /note="Organ: Hip; Vector: pBluescript; Site 1: EcoRI;
                     Library constructed by Dr. Marian Young and Dr. Pamela
                     Gehron Robey (NIDCR)"

ORIGIN

Query Match   60.1%;   Score 516;   DB 9;   Length 522;
Best Local Similarity 100.0%; Pred. No. 1.9e-257;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   327 AGAAAAACTTCAGCTCAATATAGTAGAGATGAGATGAAAATGCAACTTTAGATGGCGG 386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7 AGAAAAACTTCAGCTCAATATAGTAGAGATGAGATGAAAATGCAACTTTAGATGGCGG 66

Qy   387 AGATGTTTATTCACAGGCAGAGAATTTTTGTGGGGCCTTTCCAAAAGGACAAATCAACG 446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 AGATGTTTATTCACAGGCAGAGAATTTTTGTGGGGCCTTTCCAAAAGGACAAATCAACG 126

Qy   447 AGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGC 506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 AGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGC 186

Qy   507 AGATGGGTTGCATTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGG 566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 AGATGGGTTGCATTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGG 246

Qy   567 GTCTAGTGAATCTGCACAGAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTA 626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 GTCTAGTGAATCTGCACAGAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTA 306

Qy   627 CGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCCAA 686
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 CGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCCAA 366

Qy   687 CAAAGGGACACGTCTTGCTGCCGCACCGAACCCCGGAAGAGTATCCGAGAAGTGCAAAGGTTTA 746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 CAAAGGGACACGTCTTGCTGCCGCACCGAACCCCGGAAGAGTATCCGAGAAGTGCAAAGGTTTA 426

Qy   747 TGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGAAAAGGTGGA 806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 TGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGAAAAGGTGGA 486

Qy   807 TGGGCTGCTCACCTGCTCGCTCAGTTTTAATTAACAA 842
          |||||||||||||||||||||||||||||||||||||
Db   487 TGGGCTGCTCACCTGCTCGCTCAGTTTTAATTAACAA 522

RESULT 2
BI763990
LOCUS   BI763990   889 bp   mRNA   linear   EST 25-SEP-2001
DEFINITION 603049879f1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5189970 5',
            mRNA sequence.
ACCESSION   BI763990
VERSION   BI763990.1 GI:15755568
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 889)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.

Qy   824 GCTCAGTTTAATTAACAAGAA 845
          |||||||||||||||||||||
Db   421 GCTCAGTTTAATTAACAAGAA 442

RESULT 4
BG977444
LOCUS   BG977444   356 bp   mRNA   linear   EST 12-JUN-2001
DEFINITION RCS-CIO161-040101-013-D07 CIO161 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG977444
VERSION   BG977444.1 GI:14380179
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 356)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE   Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-CIO161-
            040101-013-D07&t3=2001-01-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 356.
FEATURES   Location/Qualifiers
     source   1..356
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CIO161"
                     /note="Organ: colon_ins; Vector: puc18; Site 1: SmaI;
                     Site 2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No.196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Query Match   40.8%; Score 350; DB 12; Length 356;
Best Local Similarity   100.0%; Pred. No. 6.8e-171;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   109 GAGGTGGACGTGCCCCGCGGAACGGCAGCACCAGCTTCACGTGGGCGTGCTGGGCAGC 168
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7 GAGGTGGACGTGCCCCGCGGAACGGCAGCACCAGCTTCACGTGGGCGTGCTGGGCAGC 66

Qy   169 AAGGTGGGGCTGCAGGTGGTGGAGCTGCCGCGGACTGCCGCAGGAGCTTCCGACGTCGCTCTTC 228
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 AAGGTGGGGCTGCAGGTGGTGGAGCTGCCGCGGACTGCCGCAGGAGCCTTCCGACGTCGCTCTTC 126

Db   604 TGGCAGATGGTTGCATTGAAGAGTT 631

RESULT 3
AY406329
LOCUS   AY406329   455 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION Homo sapiens DDAH1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406329
VERSION   AY406329.1 GI:39762303
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 455)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 455)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..455
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene   <1..>455
                     /gene="DDAH1"
                     /locus_tag="HCM2521"
ORIGIN

Query Match   51.5%; Score 442; DB 29; Length 455;
Best Local Similarity   100.0%; Pred. No. 7.3e-219;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   404 GCAGAGAATTTTTGTGGGCCTTTCAAAAGGACAAATCAACCAGAGTGCTGAAATCTTGG 463
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCAGAGAATTTTTGTGGGCCTTTCAAAAGGACAAATCAACCAGAGTGCTGAAATCTTGG 60

Qy   464 CTGATACTTTTAAGGACTATGCAGTCTCCACAGTTGCCAGTGGCAGATGGTTGCATTTGA 523
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   61 CTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGTTGCATTTGA 120

Qy   524 AGAGTTCTGCACATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCAC 583
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGAGTTCTGCACATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCAC 180

Qy   584 AGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGTACGACAAACTCACTGTGC 643
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   181 AGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGC 240

Qy   644 CTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGC 703
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   241 CTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGC 300

Qy   704 TGCACCGAACCCGGAAGAGTATCCGGAAGAGTGCAAAGGTTTATGGAAACTGAAGGACC 763
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGCACCGAACCCGGAAGAGTATCCGGAAGAGTGCAAAGGTTTATGGAAACTGAAGGACC 360

Qy   764 ATATGCTGATCCCGTGAGCATGTCGAACTGGAAAGGTGGATGGGCTGCTCACCTGCT 823

```
Qy    229  GTGGAGGACGTGCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCG  288
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127  GTGGAGGACGTGCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCG  186

Qy    289  AGCCGGGAAGGAGGTTGACATGATGAAAGAGCATTAGAAAAACTTCAGCTCAATATA  348
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187  AGCCGGGAAGGAGGTTGACATGATGAAAGAGCATTAGAAAAACTTCAGCTCAATATA  246

Qy    349  GTAGAGATGAAAGATGAAAAATGCAACTTAGATGGGGGAGATGTTTTATTCACAGGCAGA  408
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247  GTAGAGATGAAAGATGAAAAATGCAACTTAGATGGGGGAGATGTTTTATTCACAGGCAGA  306

Qy    409  GAATTTTTGTGGGCCTTTCAAAAAGACAAATCAACGAGGTGCTGAAAT  458
                ||||||||||||||||||||||||||||||||||||||||||||||||
Db    307  GAATTTTTGTGGGCCTTTCAAAAAGACAAATCAACGAGGTGCTGAAAT  356
```

RESULT 5
AY406330
LOCUS   AY406330   455 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION   Pan troglodytes DDAH1 gene, VIRTUAL TRANSCRIPT, partial sequence,
             genomic survey sequence.
ACCESSION   AY406330
VERSION   AY406330.1 GI:39762304
KEYWORDS   GSS.
SOURCE   Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 455)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
             Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
             Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
             Adams,M.D. and Cargill,M.
TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
             gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE   2 (bases 1 to 455)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
             Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
             Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
             Adams,M.D. and Cargill,M.
TITLE   Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
             Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
             them based on alignment.
FEATURES   Location/Qualifiers
     source   1..455
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene   <1..>455
                     /gene="DDAH1"
                     /locus_tag="HCM2521"
ORIGIN
Query Match   39.6%; Score 340; DB 29; Length 455;
Best Local Similarity   99.5%; Pred. No. 1.2e-165;
Matches   440; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy    404  GCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGG  463
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  GCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGG  60

Qy    464  CTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGCCAGTGGGTGCATTTGA  523
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGCCAGTGGGTGCATTTGA  120

Qy    524  AGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCCAGTGAATCTGCAC  583
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCCAGTGAATCTGCAC  180
```

```
Qy    584  AGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGC  643
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGC  240

Qy    644  CTGATGACATAGCAGCAAACTGTATATATCCAATTATATCCCAACAAAGGGCACGTCTTGC  703
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  CTGATGACGTAGCAGCAAACTGTATATATCCCAATTATATCCCAACAACAAAGGGCACGTCTTGC  300

Qy    704  TGCACCGAACCCCGAAGAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACC  763
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  TGCACCGAACCCCGAAGAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACC  360

Qy    764  ATATGCTGATCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCT  823
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  ATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCT  420

Qy    824  GCTCAGTTTAATTAACAAGAA  845
                |||||||||||||||||||||
Db    421  GCTCAGTTTAATTAACAAGAA  442
```

RESULT 6
AW514143/c
LOCUS   AW514143   624 bp   mRNA   linear   EST 03-MAR-2000
DEFINITION   hd24a02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
             IMAGE:2910410 3' similar to TR:O94760 O94760 NG,NG-DIMETHYLARGININE
             DIMETHYLAMINOHYDROLASE ;, mRNA sequence.
ACCESSION   AW514143
VERSION   AW514143.1 GI:7152225
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 624)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Seq primer: -40UP from Gibco
             High quality sequence stop: 346.
FEATURES   Location/Qualifiers
     source   1..624
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2910410"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
ORIGIN
Query Match   37.6%; Score 323; DB 10; Length 624;
Best Local Similarity 100.0%; Pred. No. 9e-157;
Matches   323; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    523  AAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCCAGTGAATCTGCA  582
```

```
Db  216  AGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAG  275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  362  ATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGCAGAATTTTTTGTGG   421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  276  ATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGCAGAATTTTTTGTGG   335
Qy  422  GCCTTTCCAAAGGACAAATCAACGAGTGCTGAAAATCTTGGCTGACTTTTAAGGACT    481
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  336  GCCTTTCCAAAGGACAAATCAACGAGTGCTGAAATCTTGGCTGACTTTTAAGGACT    395
Qy  482  ATGCAGTCTCCACAGTGCCAGTGGCAGAGGTTGCATTGAAGAGTTTCGCAGCATGG    541
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  396  ATGCAGTCTCCACAGTGCCAGTGGCAGAGGTTGCATTGAAGAGTTTCTGCAGCATGG   455
Qy  542  CTGGCCTAACCTGATCGCAATTGGGTCTAGTGAATTGTGCACAGAAGGCCCTAAGATCA  601
              ||||| ||||||||||||| |||||||||||||||| |||||||||||||||||||||
Db  456  CTGGGCCTAACCTGATCGCGATTGGGTCTAGTGAATCTGCACAGAAGGCCCTAAGATCA  515
Qy  602  TGCAACAGATGAGTGACCAC  621
              ||||||||||||||||||||
Db  516  TGCAACAGATGAGTGACCAC  535
```

RESULT 8
BF818963/c
LOCUS   BF818963   308 bp   mRNA   linear   EST 13-JAN-2001
DEFINITION RC5-CI0160-221200-012-A08 CI0160 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF818963
VERSION   BF818963.1 GI:12156591
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 308)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE   Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPSP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/getheml2.pl?t1=RC5&t2=RC5-CI0160-
            221200-012-A08&t3=2000-12-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 308.
FEATURES   Location/Qualifiers
     source   1..308
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CI0160"
                     /note="Organ: colon; colon_ins; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)

```
Db  477  AAGAGTTTCTGCAGCATGGCATGGGCCTAACCTGATCGCGCAATTGGGTCTAGTGAATCTGCA  418
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  583  CAGAGGCCCTTAAGATCACATGCAACAGATGAGTGACCACCACCGCTACGACAAACTCACTGTG  642
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  417  CAGAGGCCCTTAAGATCACATGCAACAGATGAGTGACCACCACCGCTACGACAAACTCACTGTG  358
Qy  643  CCTGATGACTAGCAGCAAACTGTATATATATCTAAATATCCCAACAAAGGGCACGTCTTG  702
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  357  CCTGATGACATAGCAGCAAACTGTATATATATCTAAATATCCCAACAAAGGGCACGTCTTG  298
Qy  703  CTGCACCGAACCCCGGAACGAGTATCCGAAAGTTCAAAGTGCAAAGTTTATGGAGAAACTGAAGGAC  762
              |||||||||||||||| |||||||||||||||||||||||||||||||| |||||||||||||||
Db  297  CTGCACCGAACCCCGGAAGAGTATCCGAAAGTTCAAAGTGCAAAGTTTATGAGAAACTGAAGGAC  238
Qy  763  CATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCCTCACCTGC  822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  237  CATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCCTCACCTGC  178
Qy  823  TGCTCAGTTTAATTAACAAGAA  845
              ||||||||||||||||||||||
Db  177  TGCTCAGTTTAATTAACAAGAA  155
```

RESULT 7
AU279491
LOCUS   AU279491   535 bp   mRNA   linear   EST 31-JUL-2003
DEFINITION AU279491 CHONS1 Homo sapiens cDNA clone CHONS1000018 5', mRNA
            sequence.
ACCESSION   AU279491
VERSION   AU279491.1 GI:2829718
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 535)
  AUTHORS   Imabayashi,H., Mori,T., Gojo,S., Kiyono,T., Sugiyama,T., Irie,R.,
            Isogai,T., Hata,J., Tomoya,Y. and Umezawa,A.
  TITLE   Redifferentiation of dedifferentiated chondrocytes and
            chondrogenesis of human bone marrow stromal cells via chondrosphere
            formation with expression profiling by large-scale cDNA analysis
  JOURNAL   Exp. Cell Res. 288 (1), 35-50 (2003)
  MEDLINE   22760698
   PUBMED   12878157
COMMENT   Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA Project; Sugiyama,T.; Wakamatsu,A.; Irie,R.;
            Umezawa,A.; Fukuma,M.; Kusakari,S.; Hata,J.; Ishii,S.; Yamamoto,J.;
            Isono,Y.; Saito,K.; Nakamura,Y.; Masuho,Y.; Nagai,K.; Isogai,T.
            HRI human cDNA project; cDNA library construction & 5'-end one
            pass sequencing: Helix Research Institute.
FEATURES   Location/Qualifiers
     source   1..535
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CHONS1000018"
                     /cell_type="chondrocytes"
                     /clone_lib="CHONS1"
                     /note="Vector: pME18SFl3"
ORIGIN

Query Match   31.4%; Score 269; DB 9; Length 535;
Best Local Similarity 99.7%; Pred. No. 1.3e-128;
Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy  302  AGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAG  361
```

ORIGIN
          Soares and M. Fatima Bonaldo. "

        Query Match   25.2%; Score 216; DB 9; Length 406;
        Best Local Similarity 100.0%; Pred. No. 5.4e-101;
        Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   530 CAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCAACAA 689
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 CAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCAACA 347

Qy   690 AGGGCACGTCTTGCTGCACCGACCCCGGAAGTATCCAGAAGTGCAAAGGTTTATGA 749
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 AGGGCACGTCTTGCTGCACCGACCCCGGAAGTATCCAGAAGTGCAAAGGTTTATGA 287

Qy   750 GAAACTGAAGGACCATATGCTGATCCCGTGAGCGAACTGCTGAACTGCTGAAAAAGGTGGATGG 809
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 GAAACTGAAGGACCATATGCTGATCCCGTGAGCGAACTGCTGAACTGCTGAAAAAGGTGGATGG 227

Qy   810 GCTGCTCACCTGCTGCTCAGTTTTAATTAACACAGAA 845
         |||||||||||||||||||||||||||||||||||||
Db   226 GCTGCTCACCTGCTGCTCAGTTTTAATTAACACAGAA 191

RESULT 10
BF818967
LOCUS   BF818967   276 bp   mRNA   linear   EST 13-JAN-2001
DEFINITION RC5-CI0160-221200-012-C11 CI0160 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF818967
VERSION   BF818967.1 GI:12155597
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 276)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE   Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE   20202663
PUBMED   10737800
COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&t2=RC5-CI0160-
            221200-012-C11&t3=2000-12-22&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 270.
FEATURES   Location/Qualifiers
     source   1..276
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CI0160"
                     /note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of

---

ORIGIN
          profiles into the pUC 18 vector. Reverse transcription of
          tissue mRNA and cDNA amplification were performed under
          low stringency conditions."

        Query Match   30.7%; Score 263; DB 10; Length 308;
        Best Local Similarity 100.0%; Pred. No. 1.6e-125;
        Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   181 CAGGTGGTGGAGACTGCCGGCGCCGACGAGAGCTTCCGGACGCTGCGTCTTCGTGGAGGACGTG 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 CAGGTGGTGGAGACTGCCGGCGCCGACGAGAGCTTCCGGACGCTGCGTCTTCGTGGAGGACGTG 249

Qy   241 GCCGTGTGTCGTGGAGACGGACGCCCTCATCACCCGACCCCGGGGGCGCCGAGCCGGAGGAAG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 GCCGTGTGTCGTGGAGACGGACGCCCTCATCACCCGACCCCGGGGGCGCCGAGCCGGAGGAAG 189

Qy   301 GAGGTTGACATGATGAAAGAAGGAGCATTAGAAAAACTTCAGTCAATATAGTAGAGATGAAA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 GAGGTTGACATGATGAAAGAAGGAGCATTAGAAAAACTTCAGTCAATATAGTAGAGATGAAA 129

Qy   361 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   128 GATGAAAATGCAACTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTTGTG 69

Qy   421 GGCCTTTCCAAAAGGACAAATCA 443
         |||||||||||||||||||||||
Db   68 GGCCTTTCCAAAAGGACAAATCA 46

RESULT 9
AI017468/c
LOCUS   AI017468   406 bp   mRNA   linear   EST 27-AUG-1998
DEFINITION ou23c07.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1627116 3' similar to TR:O08557 O08557
            N-G,N-G-DIMETHYLARGININE DIMETHYLAMINOHYDROLASE. ;, mRNA sequence.
ACCESSION   AI017468
VERSION   AI017468.1 GI:3231804
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 406)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 483   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES   Location/Qualifiers
     source   1..406
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1627116"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1:Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento

tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match   18.5%; Score 159; DB 10; Length 276;
Best Local Similarity 100.0%; Pred. No. 2.8e-71;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   220 TGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGACGGCCCTCATCACCCGACCC 279
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   118 TGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGACGGCCCTCATCACCCGACCC 177

Qy   280 GGGGCGCCGAGCCCGAGGAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAG 339
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   178 GGGGCGCCGAGCCCGAGGAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAG 237

Qy   340 CTCAATAGTAGAGATGAAGATGAAAATGCAACTTTA 378
          ||||||||||||||||||||||||||||||||||||
Db   238 CTCAATAGTAGAGATGAAGATGAAAATGCAACTTTA 276

RESULT 11
BG164988
LOCUS   BG164988   1038 bp   mRNA   linear   EST 06-FEB-2001
DEFINITION 602343844F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4454050 5',
            mRNA sequence.
ACCESSION   BG164988
VERSION   BG164988.1   GI:12671691
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1038)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10245 row: g column: 11
            High quality sequence stop: 328.
FEATURES   Location/Qualifiers
     source   1..1038
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4454050"
                     /tissue_type="hypernephroma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_89"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."

ORIGIN

Query Match   16.1%; Score 138; DB 12; Length 1038;
Best Local Similarity 100.0%; Pred. No. 3e-60;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   691 GGGGCACGTCTTCTGCTGCACCGACCCCGGAAGAGTATCCAGAAAGTGCAAAGGTTTATGAG 750
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGGCACGTCTTCTGCTGCACCGACCCCGGAAGAGTATCCAGAAAGTGCAAAGGTTTATGAG 60

Qy   751 AAACTGAAGGAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAGGTGGATGGG 810

Db   61 AAACTGAAGGAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAGGTGGATGGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   811 CTGCTCACCTGCTGCTCA 828
          ||||||||||||||||||
Db   121 CTGCTCACCTGCTGCTCA 138

RESULT 12
AA376335
LOCUS   AA376335   312 bp   mRNA   linear   EST 21-APR-1997
DEFINITION EST89005 HSC172 cells II Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION   AA376335
VERSION   AA376335.1   GI:2028887
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 312)
AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,H.W., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
TITLE   Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE   96026280
PUBMED   7566098
COMMENT   Other_ESTs: THC99618
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research.
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES   Location/Qualifiers
     source   1..312
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="ATCC (inhost):181025"
                     /db_xref="taxon:9606"
                     /cell_type="fibroblast"
                     /cell_line="HSC172 (60PDL)"
                     /dev_stage="fetal"
                     /clone_lib="HSC172 cells II"
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"

ORIGIN

Query Match   13.4%; Score 115; DB 9; Length 312;
Best Local Similarity 100.0%; Pred. No. 2.5e-48;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   303 GGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAGA 362

```
Db   161 GGTTGACATGATGATGAAAGAAGCATTAGAAGAAGACTTCAGCTCCATATAGTAGAGATGAAAGA 220
             ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   363 TGAAAAATGCAACTTAGATGGCGGAGATGTTTTATTCCACAGGCAGAGAATTTTT 417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221 TGAAAAATGCAACTTAGATGGCGGAGATGTTTTATTCCACAGGCAGAGAATTTTT 275

RESULT 13
AQ742098/c
LOCUS       AQ742098      700 bp    DNA    linear    GSS 16-JUL-1999
DEFINITION  HS_5566_B2_D01_SP6 RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1142 Col=2 Row=H, genomic survey sequence.
ACCESSION   AQ742098
VERSION     AQ742098.1  GI:5519620
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 700)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE       Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE     99380589
PUBMED      10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1142 row: H column: 2
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 700.
FEATURES             Location/Qualifiers
     source          1..700
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=1142 Col=2 Row=H"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN
Query Match   12.4%;   Score 106;   DB 28;   Length 700;
Best Local Similarity  100.0%;   Pred. No. 1.4e-43;
Matches 106;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   740 AGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCTGAACTGGAA 799
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   573 AGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCTGAACTGGAA 514
Qy   800 AGGTGGATGGGCGGCTCACCTGCTGCCCAGTTTAATTAACAAGAA 845
             |||||||||||||||||||||||||||||||||||||||||||||
Db   513 AGGTGGATGGGCGGCTCACCTGCTGCCCAGTTTAATTAACAAGAA 468
```

```
RESULT 14
BF933706
LOCUS       BF933706      167 bp    mRNA    linear    EST 22-JAN-2001
DEFINITION  CM2-NT0170-211200-647-e08 NT0170 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF933706
VERSION     BF933706.1  GI:11351030
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 167)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongenee,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     20202663
PUBMED      10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM2&t2=CM2-NT0170-
            211200-647-e08&t3=2000-12-21&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 27
            High quality sequence stop: 166.
FEATURES             Location/Qualifiers
     source          1..167
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="NT0170"
                     /note="Organ: nervous tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-Library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN
Query Match   6.9%;   Score 59;   DB 10;   Length 167;
Best Local Similarity  100.0%;   Pred. No. 3.7e-19;
Matches 59;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   739 AAGGTTTATGAGAAACTGAAGGACCATATGCTGAGCCCGTGAGCATGTCTGAACTGGA 797
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    55 AAGGTTTATGAGAAACTGAAGGACCATATGCTGAGCCCGTGAGCATGTCTGAACTGGA 113

RESULT 15
BY233360
LOCUS       BY233360      357 bp    mRNA    linear    EST 10-DEC-2002
DEFINITION  BY233360 RIKEN full-length enriched, adult inner ear Mus musculus
            cDNA clone F930030D22 5', mRNA sequence.
ACCESSION   BY233360
VERSION     BY233360.1  GI:26414470
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
```

                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="C57BL/6J"
                  /db_xref="taxon:10090"
                  /clone="F930030D22"
                  /tissue_type="inner ear"
                  /dev_stage="adult"
                  /clone_lib="RIKEN full-length enriched, adult inner ear"

ORIGIN

  Query Match   6.9%; Score 59; DB 13; Length 357;
  Best Local Similarity 100.0%; Pred. No. 4.1e-19;
  Matches   59; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

  Qy   1 ATGGCCGGCCTCGGCCACCCCTCCGCCTTCGGCCGGGGCCACCCACGCCGTGGTGCGGGC 59
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   134 ATGGCCGGCCTCGGCCACCCCTCCGCCTTCGGCCGGGGCCACCCACGCCGTGGTGCGGGC 192

Search completed: June 7, 2004, 23:11:51
Job time : 2572 secs

```
                 GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      June  7, 2004, 21:17:45 ; Search time 96 Seconds
                                 (without alignments)
                               4959.874 Million cell updates/sec

Title:       US-09-889-733B-1
Perfect score: 858
Sequence:    1 atggccggcctcggccaccc........acaagaaggtagactcctga 858

Scoring table: OLIGO_NUC
             Gapop 60.0 , Gapext 60.0

Searched:    682709 seqs, 27747446 residues

Word size :   0

Total number of hits satisfying chosen parameters:    1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   Issued Patents NA:*
             1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
             2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
             3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
             4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
             5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
             6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES

          %
Result       Query
No.   Score Match Length DB  ID                        Description
----  ----- ----- ------ --  --                        -----------
  1     19   2.2   7263   4   US-09-620-312D-55         Sequence 55, Appl
  2     18   2.1   1014   4   US-09-252-991A-9227       Sequence 9227, Ap
  3     18   2.1   1588   3   US-08-617-785-7           Sequence 7, Appli
  4     18   2.1   1588   4   US-09-817-464-7           Sequence 7, Appli
  5     18   2.1   1728   4   US-09-134-000C-1052       Sequence 1052, Ap
  6     18   2.1   2745   3   US-08-617-785-11          Sequence 11, Appl
  7     18   2.1   2745   4   US-09-817-464-11          Sequence 11, Appl
  8     18   2.1   2766   3   US-08-617-785-13          Sequence 13, Appl
  9     18   2.1   2766   4   US-09-817-464-13          Sequence 13, Appl
 10     18   2.1   2997   1   US-08-453-862-1           Sequence 1, Appl
 11     18   2.1   2997   2   US-08-452-734A-1          Sequence 1, Appli
 12     18   2.1   2997   3   US-08-176-401B-1          Sequence 1, Appli
 13     18   2.1   2997   5   PCT-US94-14989-1          Sequence 1, Appli
 14     18   2.1   3021   4   US-09-016-434-1118        Sequence 1118, Ap
 15     18   2.1   3804   3   US-08-617-785-3           Sequence 3, Appli
 16     18   2.1   3804   4   US-09-817-464-3           Sequence 3, Appli
 17     18   2.1 1664976   4  US-08-916-421B-1          Sequence 1, Appli
 18     18   2.1 1830121   4  US-09-557-884-1           Sequence 1, Appli
 19     18   2.1 1830121   4  US-09-643-990A-1          Sequence 1, Appli
 20     18   2.1 4403765   3  US-09-103-840A-2          Sequence 2, Appli
 21     18   2.1 4411529   3  US-09-103-840A-1          Sequence 1, Appli
 22     17   2.0    243    4  US-09-543-681A-3669       Sequence 3669, Ap
 23     17   2.0    515    4  US-09-621-976-12843       Sequence 12843, A
 24     17   2.0    524    4  US-09-621-976-1137        Sequence 1137, Ap
 25     17   2.0    528    4  US-09-134-001C-1910       Sequence 1910, Ap
 26     17   2.0    579    4  US-09-449-285A-13         Sequence 13, Appl
 27     17   2.0    586    3  US-09-328-111-588         Sequence 588, App
```

```
 28     17   2.0   696   4  US-09-252-991A-2735    Sequence 2735, Ap
 29     17   2.0   792   4  US-09-252-991A-7444    Sequence 7444, App
 30     17   2.0   807   4  US-09-252-991A-164     Sequence 164, App
 31     17   2.0   891   4  US-09-252-991A-176     Sequence 176, App
 32     17   2.0   948   4  US-09-252-991A-172     Sequence 172, App
 33     17   2.0   960   4  US-09-252-991A-7256    Sequence 7256, Ap
 34     17   2.0  1007   4  US-09-247-155-81       Sequence 81, Appl
 35     17   2.0  1069   4  US-09-372-422A-7       Sequence 7, Appli
 36     17   2.0  1088   4  US-09-148-545-132      Sequence 132, App
 37     17   2.0  1143   4  US-09-148-545-79       Sequence 79, Appl
 38     17   2.0  1168   4  US-09-484-970B-89      Sequence 89, Appl
 39     17   2.0  1333   4  US-09-372-422A-9       Sequence 9, Appli
 40     17   2.0  1354   4  US-09-620-312D-308     Sequence 308, App
 41     17   2.0  1611   4  US-09-252-991A-183     Sequence 183, App
 42     17   2.0  2338   3  US-09-232-200-66       Sequence 66, Appl
 43     17   2.0  2338   4  US-09-232-197-66       Sequence 66, Appl
 44     17   2.0  2338   4  US-09-232-201-66       Sequence 66, Appl
 45     17   2.0  2338   4  US-09-232-195-66       Sequence 66, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-620-312D-55
; Sequence 55, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Liu, Chenghua
;  APPLICANT: Asundi, Vinod
;  APPLICANT: Zhang, Jie
;  APPLICANT: Ren, Feiyan
;  APPLICANT: Chen, Rui-hong
;  APPLICANT: Zhao, Qing A.
;  APPLICANT: Wehrman, Tom
;  APPLICANT: Xue, Aidong J.
;  APPLICANT: Yang, Yonghong
;  APPLICANT: Wang, Jian-Rui
;  APPLICANT: Zhou, Ping
;  APPLICANT: Ma, Yunqing
;  APPLICANT: Wang, Dunrui
;  APPLICANT: Wang, Zhiwei
;  APPLICANT: John Tillinghast
;  APPLICANT: Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
;  TITLE OF INVENTION: Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_Fl_genes Version 1.0
; SEQ ID NO 55
;  LENGTH: 7263
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (159)..(5057)
US-09-620-312D-55

  Query Match             2.2%;  Score 19;  DB 4;  Length 7263;
  Best Local Similarity   100.0%;  Pred. No. 7.3;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       178 CTGCAGGTGGTGGAGCTGC 196
             |||||||||||||||||||
Db      1251 CTGCAGGTGGTGGAGCTGC 1269
```

RESULT 2
US-09-252-991A-9227
; Sequence 9227, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9227
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9227

Query Match   2.1%;   Score 18;   DB 4;   Length 1014;
Best Local Similarity   100.0%;   Pred. No. 23;
Matches 18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   193 CTGCCGGCCGACGAGAGC 210
          ||||||||||||||||||
Db   811 CTGCCGGCCGACGAGAGC 828

RESULT 3
US-08-617-785-7
; Sequence 7, Application US/08617785E
; Patent No. 6228610
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/08/617,785E
; CURRENT FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: PCT/EP94/02991
; EARLIER FILING DATE: 1994-09-07
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1588
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1447)
US-08-617-785-7

Query Match   2.1%;   Score 18;   DB 3;   Length 1588;
Best Local Similarity   100.0%;   Pred. No. 23;
Matches 18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   332 AACTTCAGCTCAATATAG 349
          |||||||||||||||||
Db   228 AACTTCAGCTCAATATAG 245

RESULT 4
US-09-817-464-7
; Sequence 7, Application US/09817464
; Patent No. 6515107
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/09/817,464
; CURRENT FILING DATE: 2001-03-26
; EARLIER APPLICATION NUMBER: US/08/617,785
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1588
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1447)
US-09-817-464-7

Query Match   2.1%;   Score 18;   DB 4;   Length 1588;
Best Local Similarity   100.0%;   Pred. No. 23;
Matches 18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   332 AACTTCAGCTCAATATAG 349
          |||||||||||||||||
Db   228 AACTTCAGCTCAATATAG 245

RESULT 5
US-09-134-000C-1052/c
; Sequence 1052, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1052
; LENGTH: 1728
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1052

Query Match   2.1%;   Score 18;   DB 4;   Length 1728;
Best Local Similarity   100.0%;   Pred. No. 23;
Matches 18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   652 ATAGCAGCAAACTGTATA 669
          |||||||||||||||||
Db   565 ATAGCAGCAAACTGTATA 548

RESULT 6
US-08-617-785-11

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    332 AACTTCAGCTCAATATAG 349
          |||||||||||||||||
Db   1502 AACTTCAGCTCAATATAG 1519
```

RESULT 8
US-08-617-785-13
; Sequence 13, Application US/08617785E
; Patent No. 6228610
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/08/617,785E
; CURRENT FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: PCT/EP94/02991
; EARLIER FILING DATE: 1994-09-07
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2766
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2766)
US-08-617-785-13

Query Match   2.1%;   Score 18;   DB 3;   Length 2766;
Best Local Similarity   100.0%;   Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    332 AACTTCAGCTCAATATAG 349
          |||||||||||||||||
Db   1502 AACTTCAGCTCAATATAG 1519
```

RESULT 9
US-09-817-464-13
; Sequence 13, Application US/09817464
; Patent No. 6515107
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/09/817,464
; CURRENT FILING DATE: 2001-03-26
; EARLIER APPLICATION NUMBER: US/08/617,785
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2766

; Sequence 11, Application US/08617785E
; Patent No. 6228610
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/08/617,785E
; CURRENT FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: PCT/EP94/02991
; EARLIER FILING DATE: 1994-09-07
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2745)
US-08-617-785-11

Query Match   2.1%;   Score 18;   DB 3;   Length 2745;
Best Local Similarity   100.0%;   Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    332 AACTTCAGCTCAATATAG 349
          |||||||||||||||||
Db   1502 AACTTCAGCTCAATATAG 1519
```

RESULT 7
US-09-817-464-11
; Sequence 11, Application US/09817464
; Patent No. 6515107
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/09/817,464
; CURRENT FILING DATE: 2001-03-26
; EARLIER APPLICATION NUMBER: US/08/617,785
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2745)
US-09-817-464-11

Query Match   2.1%;   Score 18;   DB 4;   Length 2745;
Best Local Similarity   100.0%;
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: (1)..(2766)
US-09-817-464-13

Query Match             2.1%;  Score 18;  DB 4;  Length 2766;
Best Local Similarity   100.0%;  Pred. No. 23;
Matches    18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        332 AACTTCAGCTCAATATAG 349
               |||||||||||||||||
Db       1502 AACTTCAGCTCAATATAG 1519


RESULT 10
US-08-453-862-1
;  Sequence 1, Application US/08453862
;  Patent No. 5738999
;  GENERAL INFORMATION:
;     APPLICANT: Segerson, Thomas P.
;     APPLICANT: Kinzie, J. Mark
;     APPLICANT: Mulvihill, Eileen R.
;     APPLICANT: Saugstad, Julie A.
;     APPLICANT: Westbrook, Gary L.
;  TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS
;  NUMBER OF SEQUENCES: 5
;  CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Townsend and Townsend and Crew LLP
;     STREET: Two Embarcadero Center, 8th Floor
;     CITY: San Francisco
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94111-3834
;  COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/453,862
;     FILING DATE: 30-MAY-1995
;     CLASSIFICATION: 436
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/176,401
;     FILING DATE: 30-DEC-1993
;  ATTORNEY/AGENT INFORMATION:
;     NAME: Parmelee, Steven W.
;     REGISTRATION NUMBER: 31,990
;     REFERENCE/DOCKET NUMBER: 13952-18-2
;  TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 206-467-9600
;     TELEFAX: 415-576-0300
;  INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 2997 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     FEATURE:
;        NAME/KEY: CDS
;        LOCATION: 133..2877
US-08-453-862-1

Query Match             2.1%;  Score 18;  DB 1;  Length 2997;
Best Local Similarity   100.0%;  Pred. No. 23;
Matches    18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        332 AACTTCAGCTCAATATAG 349
               |||||||||||||||||
Db       1634 AACTTCAGCTCAATATAG 1651
```

```
RESULT 11
US-08-452-734A-1
;  Sequence 1, Application US/08452734A
;  Patent No. 5831047
;  GENERAL INFORMATION:
;     APPLICANT: Segerson, Thomas P.
;     APPLICANT: Kinzie, J. Mark
;     APPLICANT: Mulvihill, Eileen R.
;     APPLICANT: Saugstad, Julie A.
;     APPLICANT: Westbrook, Gary L.
;  TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS
;  NUMBER OF SEQUENCES: 5
;  CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Townsend and Townsend and Crew LLP
;     STREET: Two Embarcadero Center, 8th Floor
;     CITY: San Francisco
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94111-3834
;  COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/452,734A
;     FILING DATE: 30-MAY-1995
;     CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/176,401
;     FILING DATE: 30-DEC-1993
;  ATTORNEY/AGENT INFORMATION:
;     NAME: Parmelee, Steven W.
;     REGISTRATION NUMBER: 31,990
;     REFERENCE/DOCKET NUMBER: 13952-18-1
;  TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 206-467-9600
;     TELEFAX: 415-575-0300
;  INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 2997 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     FEATURE:
;        NAME/KEY: CDS
;        LOCATION: 133..2877
US-08-452-734A-1

Query Match             2.1%;  Score 18;  DB 2;  Length 2997;
Best Local Similarity   100.0%;  Pred. No. 23;
Matches    18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        332 AACTTCAGCTCAATATAG 349
               |||||||||||||||||
Db       1634 AACTTCAGCTCAATATAG 1651


RESULT 12
US-08-176-401B-1
;  Sequence 1, Application US/08176401B
;  Patent No. 6274330
;  GENERAL INFORMATION:
;     APPLICANT: Segerson, Thomas P.
;     APPLICANT: Kinzie, J. Mark
;     APPLICANT: Mulvihill, Eileen R.
;     APPLICANT: Saugstad, Julie A.
;     APPLICANT: Westbrook, Gary L.
;  TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS
;  NUMBER OF SEQUENCES: 5
```

```
;    CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Townsend and Townsend and Crew LLP
;     STREET: Two Embarcadero Center, 8th Floor
;     CITY: San Francisco
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94111-3834
;    COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/176,401B
;     FILING DATE: 30-DECEMBER-1993
;     CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;     NAME: Parmelee, Steven W.
;     REGISTRATION NUMBER: 31,990
;     REFERENCE/DOCKET NUMBER: 13952-18-2
;    TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 206-467-9600
;     TELEFAX: 415-576-0300
;    INFORMATION FOR SEQ ID NO: 1:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 2997 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     FEATURE:
;       NAME/KEY: CDS
;       LOCATION: 133..2877
US-08-176-401B-1

Query Match           2.1%;  Score 18;  DB 3;  Length 2997;
Best Local Similarity  100.0%;  Pred. No. 23;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    332 AACTTCAGCTCAATATAG 349
          |||||||||||||||||
Db   1634 AACTTCAGCTCAATATAG 1651

RESULT 13
PCT-US94-14989-1
; Sequence 1, Application PC/TUS9414989
; GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS
;    NUMBER OF SEQUENCES: 3
;    COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US94/14989
;     FILING DATE:
;     CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/176,401
;     FILING DATE: 30-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;     REGISTRATION NUMBER: 31,990
;     NAME: Parmelee, Steven W.
;     REFERENCE/DOCKET NUMBER: 13952-18PC
;    TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (206) 467-9600
;     TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
   LENGTH: 2997 base pairs
```

```
;    TYPE: nucleic acid
;    STRANDEDNESS: single
;    TOPOLOGY: linear
;    MOLECULE TYPE: cDNA
;    FEATURE:
;       NAME/KEY: CDS
;       LOCATION: 133..2877
PCT-US94-14989-1

Query Match           2.1%;  Score 18;  DB 5;  Length 2997;
Best Local Similarity  100.0%;  Pred. No. 23;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    332 AACTTCAGCTCAATATAG 349
          |||||||||||||||||
Db   1634 AACTTCAGCTCAATATAG 1651

RESULT 14
US-09-016-434-1118
; Sequence 1118, Application US/09016434
; Patent No.6500938
; GENERAL INFORMATION:
;    APPLICANT: Janice Au-Young
;    APPLICANT: Jeffrey J. Seilhamer
;    TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION: PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES: 1490
;    CORRESPONDENCE ADDRESS:
;     ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;     STREET: 3174 PORTER DRIVE
;     CITY: PALO ALTO
;     STATE: CALIFORNIA
;     COUNTRY: USA
;     ZIP: 94304
;    COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/016,434
;     FILING DATE: HEREWITH
;     CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;     CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;     NAME: Zeller, Karen J.
;     REGISTRATION NUMBER: 37,071
;     REFERENCE/DOCKET NUMBER: PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (650) 855-0555
;     TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1118:
SEQUENCE CHARACTERISTICS:
;     LENGTH: 3021 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     IMMEDIATE SOURCE:
;      LIBRARY: GENBANK
;      CLONE: g1370110
US-09-016-434-1118

Query Match           2.1%;  Score 18;  DB 4;  Length 3021;
Best Local Similarity  100.0%;  Pred. No. 23;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    332 AACTTCAGCTCAATATAG 349
          |||||||||||||||||
Db   1740 AACTTCAGCTCAATATAG 1757
```

RESULT 15
US-08-617-785-3
; Sequence 3, Application US/08617785E
; Patent No. 6228610
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/08/617,785E
; CURRENT FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: PCT/EP94/02991
; EARLIER FILING DATE: 1994-09-07
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2604)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3325)..(3495)
; OTHER INFORMATION: nucleotides designated as n could be a or g or c
; OTHER INFORMATION: or t/u
US-08-617-785-3

Query Match   2.1%;   Score 18;   DB 3;   Length 3804;
Best Local Similarity   100.0%;   Pred. No. 23;
Matches   18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   332 AACTTCAGCTCAATATAG 349
          |||||||||||||||||
Db   1337 AACTTCAGCTCAATATAG 1354

Search completed: June 7, 2004, 23:13:46
Job time : 105 secs

```
            GenCore version 5.1.6
       Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    June 7, 2004, 21:25:06 ; Search time 417 Seconds
                                     (without alignments)
                          9386.553 Million cell updates/sec

Title:           US-09-889-733B-1
Perfect score:   858
Sequence:        1 atggccggctcggccaccc........acaagaaggtagactcctga 858

Scoring table:   OLIGO_NUC
                 Gapop 60.0 , Gapext 60.0

Searched:        2995936 seqs, 2280998010 residues

Word size :      0

Total number of hits satisfying chosen parameters:     5991872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :       Published_Applications_NA:*
    1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
    2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
    3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
    4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
    5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
    6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
    7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
    8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
    9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
   10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
   11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
   12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
   13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
   14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
   15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
   16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
   17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
   18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
   19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	772	90.0	1633	17	US-10-450-826-9	Sequence 9, Appli
2	772	90.0	4208	17	US-10-468-125-22	Sequence 22, Appl
3	610	71.1	3651	9	US-09-925-300-512	Sequence 512, App
4	585	68.2	3778	15	US-10-252-157-328	Sequence 328, App
5	148	17.2	83946	17	US-10-450-826-8	Sequence 8, Appli
6	108	12.6	3189	10	US-09-814-353-20553	Sequence 20553, A
7	60	7.0	60	10	US-09-908-975-9030	Sequence 9030, Ap
8	54	6.3	3008	15	US-09-205-219-83	Sequence 83, Appl
9	33	3.8	332	9	US-09-783-599-5309	Sequence 5309, Ap
10	23	2.7	2523	15	US-10-246-785-16	Sequence 16, Appl
11	23	2.7	2771	10	US-09-361-652-4	Sequence 4, Appli
12	23	2.7	2771	10	US-09-927-315-4	Sequence 4, Appli
13	23	2.7	2771	15	US-10-190-417-4	Sequence 4, Appli
14	20	2.3	21470	9	US-09-764-847-1157	Sequence 1157, Ap
15	20	2.3	21470	15	US-10-092-154-1157	Sequence 1157, Ap
16 c	20	2.3	1691139	15	US-10-067-514-1	Sequence 1, Appli
17 c	20	2.3	1691139	16	US-10-419-723-1	Sequence 1, Appli
18 c	19	2.2	368	9	US-09-764-847-1285	Sequence 1285, Ap
19 c	19	2.2	368	15	US-10-092-154-1285	Sequence 1285, Ap
20 c	19	2.2	783	15	US-10-156-761-1744	Sequence 1744, Ap
21 c	19	2.2	954	13	US-10-282-122A-9827	Sequence 9827, Ap
22	19	2.2	1065	16	US-10-369-493-43191	Sequence 43191, A
23	19	2.2	2033	13	US-10-425-114-11394	Sequence 11394, A
24	19	2.2	2056	13	US-10-424-599-117586	Sequence 117586,
25 c	19	2.2	4371	10	US-09-814-353-21999	Sequence 21999, A
26	19	2.2	4412	10	US-09-927-827-43	Sequence 43, Appl
27	19	2.2	7263	15	US-10-037-270-55	Sequence 55, Appl
28	19	2.2	7263	16	US-10-117-722-55	Sequence 55, Appl
29	19	2.2	8148	10	US-09-764-891-7783	Sequence 7783, Ap
30	19	2.2	47753	13	US-10-087-192-1474	Sequence 1474, Ap
31 c	19	2.2	3186778	16	US-10-027-632-174961	Sequence 174961,
32 c	19	2.2	3186778	15	US-10-027-632-174961	Sequence 174961,
33	19	2.2	9025608	14	US-10-156-761-1	Sequence 1, Appli
34 c	18	2.1	269	9	US-10-040-739-1107	Sequence 1107, Ap
35	18	2.1	280	9	US-09-867-550-611	Sequence 611, App
36 c	18	2.1	435	13	US-09-920-300A-247	Sequence 247, App
37	18	2.1	435	14	US-10-240-425-429	Sequence 429, App
38	18	2.1	435	15	US-10-033-528-247	Sequence 247, App
39	18	2.1	442	9	US-09-099-926-247	Sequence 247, App
40	18	2.1	442	9	US-09-864-761-4857	Sequence 4857, Ap
41 c	18	2.1	480	16	US-10-027-632-189822	Sequence 189822,
42 c	18	2.1	480	9	US-10-027-632-189822	Sequence 189822,
43	18	2.1	496	9	US-09-960-253-125	Sequence 125, App
44 c	18	2.1	527	13	US-10-027-632-34714	Sequence 34714, A
45 c	18	2.1	527	16	US-10-027-632-34714	Sequence 34714, A

ALIGNMENTS

```
RESULT 1
US-10-450-826-9
; Sequence 9, Application US/10450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
;    APPLICANT: Ji, Darren
;    APPLICANT: Axelrod, Douglas W.
;    APPLICANT: Cook, Jonathon S.
;    APPLICANT: Jaiswal, Neelam
;    APPLICANT: Eistein, Richard
;    APPLICANT: Houghton, Adam
;    APPLICANT: Mertz, Lawrence
;    TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiatic
;    FILE REFERENCE: 044921-5039-WO
;    CURRENT APPLICATION NUMBER: US/10/450,826
;    CURRENT FILING DATE: 2003-06-18
;    PRIOR APPLICATION NUMBER: US 60/255,882
;    PRIOR FILING DATE: 2000-12-18
;    PRIOR APPLICATION NUMBER: US 60/285,691
;    PRIOR FILING DATE: 2001-04-24
;    NUMBER OF SEQ ID NOS: 149
;    SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;    LENGTH: 1633
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    OTHER INFORMATION: Genbank Accession No. AB001915
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (1)..(1633)
;    OTHER INFORMATION: n = a or c or g or t
US-10-450-826-9

Query Match             90.0%;  Score 772;  DB 17;  Length 1633;
Best Local Similarity   99.9%;            Pred. No. 0;
Matches  822;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

APPLICANT: RICHARDSON, Thomas W.
APPLICANT: RING, Huijun Z.
APPLICANT: SANJANWALA, Madhusudan
APPLICANT: SWARNAKAR, Anita
APPLICANT: CHAWLA, Narinder K.
APPLICANT: WARREN, Bridget A.
APPLICANT: XU, Yuming
APPLICANT: YUE, Henry
APPLICANT: ZEBARJADIAN, Yeganeh
TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
FILE REFERENCE: PI-0363 USN
CURRENT APPLICATION NUMBER: US/10/468,125
CURRENT FILING DATE: 2003-08-15
PRIOR APPLICATION NUMBER: PCT/US02/04918
PRIOR FILING DATE: 2002-02-14
PRIOR APPLICATION NUMBER: US 60/269,643
PRIOR FILING DATE: 2001-02-16
PRIOR APPLICATION NUMBER: US 60/271,332
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: US 60/276,767
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: US 60/282,077
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 60/285,447
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 60/287,060
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/288,543
PRIOR FILING DATE: 2001-05-03
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PERL Program
SEQ ID NO 22
LENGTH: 4208
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 7671089CB1
US-10-468-125-22

```
Query Match             90.0%;   Score 772;  DB 17;  Length 4208;
Best Local Similarity   99.9%;   Pred. No. 0;
Matches   822;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy     23 CCGCCTTCGGCGGCGGGCCACCCACGCCGTGGTGCGGGCGCTACCCGAGTCGCTCTGCCAGC 82
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 CCGCCTTCGGCGGCGGGCCACCCACGCCGTGGTGCGGGCGCTACCCGAGTCGCTCTGCCAGC 82

Qy    345 CCGCCTTCGGCGGCGGGCCACCCACGCCGTGGTGCGGGCGCTACCCGAGTCGCTCGGCCAGC 404
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 CCGCCTTCGGCGGCGGGCCACCCACGCCGTGGTGCGGGCGCTACCCGAGTCGCTCGGCCAGC 449

Qy     83 ACGCGCTGAGAAGCGCCAAGGCGCGAGGAGGTGGACGTCGCCGCCGCGGAACGGCAGCACC 142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    405 ACGCGCTGAGAAGCGCCAAGGCGCGAGGAGGTGGACGTCGCCGCCGCGGAACGGCAGCACC 464

Qy     83 ACGCGCTGAGAAGCGCCAAGGCGCGAGGAGGTGGACGTCGCCGCCGCGGAACGGCAGCACC 142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 ACGCGCTGAGAAGCGCCAAGGCGCGAGGAGGTGGACGTCGCCGCCGCGGAACGGCAGCACC 509

Qy    143 AGCTCTACGTGGGCGTGCTGGGCAGCAAGCTCGGGCTGCAGGTGGTGGAGCTGCCGGCCG 202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450 ACGCGCTGAGAAGCGCCAAGGCGCGAGGAGGTGGACGTCGCCGCCGCGGAACGGCAGCACC 569

Qy    465 AGCTCTACGTGGGCGTGCTGGGCAGCAAGCTCGGGCTGCAGGTGGTGGAGCTGCCGGCCG 524
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 AGCTCTACGTGGGCGTGCTGGGCAGCAAGCTCGGGCTGCAGGTGGTGGAGCTGCCGGCCG 202

Qy    203 ACGAGAGCCTTCCGGACTGCGCTTCGTTCGTCGGAGGACGTGGCCGTGGTGTGCGGAGGAGACGG 262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510 AGCTCTACGTGGGCGTGCTGGGCAGCAAGCTCGGGCTGCAGGTGGTGGAGCTGCCGGCCG 569

Qy    525 ACGAGAGCCTTCCGGACTGCGTTCGTTCGTCGGAGGACGTGGCCGTGGTGTGCGGAGGAGACGG 584
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 ACGAGAGCCTTCCGGACTGCGCTTCGTTCGTCGGAGGACGTGGCCGTGGTGTGCGGAGGAGACGG 262

Qy    263 CCCTCATCACCCGACCCGGGGCGCCGAGCCCGGAGGAAGGAGTTGACATGATGATGAAAGAAG 322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    570 ACGAGAGCCTTCCGGACTGCGTTCGTTCGTCGGAGGACGTGGCCGTGGTGTGCGGAGGAGACGG 629

Qy    585 CCCTCATCACCCGACCCGGGGCGCCGAGCCCGGAGGAAGGAGTTGACATGATGATGAAAGAAG 644
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 CCCTCATCACCCGACCCGGGGCGCCGAGCCCGGAGGAAGGAGTTGACATGATGATGAAAGAAG 322

Qy    323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATG 382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    630 CCCTCATCACCCGACCCGGGGCGCCGAGCCCGGAGGAAGGAGTTGACATGATGATGAAAGAAG 689

Qy    645 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATG 704
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATG 382

Qy    383 GCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC 442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    690 CATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATG 749

Qy    705 GCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC 764
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 GCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC 442

Qy    443 AACGAGGTGCTGAAATCTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAG 502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    750 GCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATC 809

Qy    765 AACGAGGTGCTGAAATCTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAG 824
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 AACGAGGTGCTGAAATCTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAG 502

Qy    503 TGCAGATGGGGTGCATTGAAGAGTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAA 562
          
Db    825 TGCAGATGGGGTGCATTGAAGAGTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAA 884

Qy    563 TTGGGTCAGTGAATCTGCACAGAAGGCCCTTAAGATCATCGCAACAGAGTGACCACC 622
          
Db    885 TTGGGTCAGTGAATCTGCACAGAAGGCCCTTAAGATCATCGCAACAGAGTGACCACC 944

Qy    623 GCTACGACAAAACTCACTGCTGATGACATAGCAGCAAACTGTATATCCTAAAATATCC 682
          
Db    945 GCTACGACAAAACTCACTGCTGATGACATAGCAGCAAACTGTATATCCTAAAATATCC 1004

Qy    683 CCAACAAAGGGCACGTCTTGCTGCACCGAACCCGAAGAGTATCCAGAAAGTGCAAAGG 742
          
Db   1005 CCAACAAAGGGCACGTCTTGCTGCACCGAACCCGAAGAGTATCCAGAAAGTGCAAAGG 1064

Qy    743 TTTATGAGAAAACTGAAGGACCATATGCTGATCCCCGTGAACTGGAAAAGG 802
          
Db   1065 TTTATGAGAAAACTGAAGGACCATATGCTGATCCCCGTGAACTGGAAAAGG 1124

Qy    803 TGGATGGGCTGCTCACCTGCTCAGTTTAATTAACAAGAA 845
          
Db   1125 TGGATGGGCTGCTCACCTGCTCAGTTTAATTAACAAGAA 1167
```

RESULT 2
US-10-468-125-22
; Sequence 22, Application US/10468125
; Publication No. US20040082061A1
; GENERAL INFORMATION:
APPLICANT: ASTROMOFF, Anna
APPLICANT: AU-YOUNG, Janice
APPLICANT: BAUGHN, Mariah R.
APPLICANT: DING, Li
APPLICANT: DUGGAN, Brendan M.
APPLICANT: FORSYTHE, Ian J.
APPLICANT: GIETZEN, Kimberly J.
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: LEE, Ernestine A.
APPLICANT: LU, Yan

```
Qy   335 TTCAGCTCAATATAGTAGAGAGATGAAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTT 394
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 TTCAGCTCAATATAGTAGAGAGATGAAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTT 181

Qy   395 TATTCACAGGCAGAGAATTTTTGTGGCCTTCAAAAGACAAATCAACAGAGGTGCTG 454
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 TATTCACAGGCAGAGAATTTTTGTGGCCTTCCAAAAGGACAAATCAACAGAGGTGCTG 241

Qy   455 AAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGCCAGATGGGT 514
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 AAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGCCAGATGGGT 301

Qy   515 TGCATTTGAAGAGTTTCTGCACGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTG 574
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 TGCATTTGAAGAGTTTCTGCACGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTCAGTG 361

Qy   575 AATCTGCACAGAAGGCCCTTAGATCATGCAACAGATGAGTGACCACCGCTACGACAAAC 634
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 AATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAAC 421

Qy   635 TCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCAACAACAAGGGC 694
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 TCACTGTGCCTGATGACATAGCAGCAAACTGTATATATATCCCAACAACAAAGGGC 481

Qy   695 ACGTCTTGCTGCACCGAACCCCGGAGAGTATCCAGAAAGTGCAAAGTTTATGAGAAAC 754
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 ACGTCTTGCTGCACCGAACCCCGGAGAGTATCCAGAAAGTGCAAAGTTTATGAGAAAC 541

Qy   755 TGAAGGACCAPATGCTGATCCCCGTGAGCATGTCTGAACTGGATGGGGCTGC 814
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   542 TGAAGGACCAATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGGGCTGC 601

Qy   815 TCACCTGCTG 824
         ||||||||||
Db   602 TCACCTGCTG 611
```

RESULT 4
US-10-252-157-328
; Sequence 328, Application US/10252157
; Publication No. US20030119640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 328
; LENGTH: 3778
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030119640A1 198087.1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1083-1106
; OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-328

Query Match   68.2%;   Score 585; DB 15; Length 3778;
Best Local Similarity 99.9%;   Pred. No. 3.9e-302;
Matches 705; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy   141 CCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCTGCAGGTGGTGGAGCTGCCGGC 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCTGCAGGTGGTGGAGCTGCCGGC 60
```

```
Db   810 AACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCCAG 869
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   503 TGGCAGATGGGTTGCATTTGAAGAGTTTCTGCACGCATGGCTGGGCCTAACCTGATCGCAA 562

Db   870 TGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAA 929
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   563 TTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACC 622

Db   930 TTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCACC 989
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   623 GCTACGACAAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATATCCTAAAATATCC 682

Db   990 GCTACGACAAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATATCCTAAAATATCC 1049
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   683 CCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAGAGTATCCAGAAAGTGCAAAGG 742

Db  1050 CCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAGAGTATCCAGAAAGTGCAAAGG 1109
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   743 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGG 802

Db  1110 TTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGG 1169
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   803 TGGATGGGGCTGCTCACCTGCTGCCTCAGTTTTAATTAACAAGAA 845

Db  1170 TGGATGGGGCTGCTCACCTGCTGCCTCAGTTTTAATTAACAAGAA 1212
```

RESULT 3
US-09-925-300-512
; Sequence 512, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 512
; LENGTH: 3651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1283)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (3641)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (3650)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-512

Query Match   71.1%;   Score 610; DB 9; Length 3651;
Best Local Similarity 100.0%;   Pred. NO. 0;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   215 CGGACTGCGTCTTCGTGAGGACGTGGCCGTGGTGTGCGAGGAGACGGCCCTCATCACCC 274
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2 CGGACTGCGTCTTCGTGAGGACGTGGCCGTGGTGTGCGAGGAGACGGCCCTCATCACCC 61

Qy   275 GACCCGGCGGCGCCGAGCCCGAGGAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAAC 334
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    62 GACCCGGCGGCGCCGAGCCCGAGGAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAAC 121
```

```
Qy   201  CGACGAGAGCCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCGCCGTGGTGTGCGAGGAGAC  260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CGACGAGAGCCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCGCCGTGGTGTGCGAGGAGAC  120

Qy   261  GGGCCCTCATCACCCGACCCGGGCGCCGAGCCCGGAGGAAGGAGGTTGACATGATGAAAGA  320
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   121  GGGCCCTCATCACCCGACCCGGGCGGCGCGAGCCCGGAGGAAGGAGGTTGACATGATGAAAGA  180

Qy   321  AGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGA  380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  AGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGA  240

Qy   381  TGGGGAGATGTTTATTCACAGGCAGGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAA  440
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TGGCGGAGATGTTTATTCACAGGCAGGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAA  300

Qy   441  T-CAACGAGGTGCTGAAATCTTGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGC  499
          |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  TCCAACGAGGTGCTGAAATCTTGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGC  360

Qy   500  CAGTGGCAGATGGGTTGCATTGGAAGAGTTTCGCAGCATGGCTGGGCCTAACCTGATCG  559
          ||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||
Db   361  CAGTGGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCG  420

Qy   560  CAATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACC  619
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CAATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACC  480

Qy   620  ACCGCTACGACAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATA  679
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  ACCGCTACGACAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATA  540

Qy   680  TCCCCAACAAAGGGCACGTCTTGCTGCACGAACCCCGGAAGAGTATCCAGAAAGTGCAA  739
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  TCCCCAACAAAGGGCACGTCTTGCTGCACGAACCCCGGAAGAGTATCCAGAAAGTGCAA  600

Qy   740  AGGTTTATGAGAAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGAAA  799
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  AGGTTTATGAGAAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGAAA  660

Qy   800  AGGTTGGATGGGCTGCTCACCTGCTCCTCAGTTTTAATTAACAAGAA  845
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   661  AGGTTGGATGGGCTGCTCACCTGCTCCTCAGTTTTAATTAACAAGAA  706
```

RESULT 5
US-10-450-826-8/c
; Sequence 8, Application US/10450826
; Publication No. US2004010181A1
; GENERAL INFORMATION:
; APPLICANT: Ji, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathon S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Eistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiati
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 83946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AL078459

US-10-450-826-8

Query Match   17.2%; Score 148; DB 17; Length 83946;
Best Local Similarity   100.0%; Pred. No. 2.6e-68;
Matches 148; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    596  AGATCATGCAACAGATGAGTGACCTACGACACCGCTACGACAAACTCACTGTGCCTGATGACATAG   655
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  18612  AGATCATGCAACAGATGAGTGACCTACGACACCGCTACGACAAACTCACTGTGCCTGATGACATAG  18553

Qy    656  CAGCAAACTGTATATATCTAAATATCCCAACAAATATCCCAACAAAGGGCAGTCTTGCTGCTGACCGAACCC   715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  18552  CAGCAAACTGTATATATCTAAATATCCCAACAAATATCCCAACAAAGGGCACGTCTTGCTGCTGACCGAACCC  18493

Qy    716  CGGAAGAGTATCCAGAAAGTGCAAAGT   743
           |||||||||||||||||||||||||||
Db  18492  CGGAAGAGTATCCAGAAAGTGCAAAGT  18465
```

RESULT 6
US-09-814-353-20553/c
; Sequence 20553, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20553
; LENGTH: 3189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3189
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20553

RESULT 7
US-09-908-975-9030
; Sequence 9030, Application US/09908975

US-09-908-975-9030

Query Match   12.6%; Score 108; DB 10; Length 3189;
Best Local Similarity   100.0%; Pred. No. 6.4e-47;
Matches 108; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    738  AAAGGTTTATGAGAAACTGAAGGACCATAATGCTGATCCCCGTGAGCATGTCTGAACTGGA   797
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3187  AAAGGTTTATGAGAAACTGAAGGACCATAATGCTGATCCCCGTGAGCATGTCTGAACTGGA  3128

Qy    798  AAAGGTGGATGGGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA   845
           |||||||||||||||||||||||||||||||||||||||||||||
Db   3127  AAAGGTGGATGGGCTGCTCACCTGCTCAGTTTTAATTAACAAGAA  3080
```

; Publication No.US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRRNT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9030
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-9030

Query Match   7.0%;   Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 3e-21;
Matches 60; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   751 AAACTGAAGGACCATATGCTGATCCCCGTGAACTGGAAAAGGTGGATGGG 810
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AAACTGAAGGACCATATGCTGATCCCCGTGAACTGGAAAAGGTGGATGGG 60

RESULT 8
US-10-205-219-83
; Sequence 83, Application US/10205219
; Publication No.US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: N-G, N-G dimethylarginine dimethylaminohydrolase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2411, 2412
; OTHER INFORMATION: n is a or g or c or t
US-10-205-219-83

Query Match   6.3%;   Score 54; DB 15; Length 3008;
Best Local Similarity 100.0%; Pred. No. 5e-18;
Matches 54; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   136 CAGCACCAGTCTCACGTGGGCGTGCTGGGCCAGCAAGCTGGGGCTGCAGGTGGTG 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 CAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGGTGGTG 620

RESULT 9

US-09-783-590-5309
; Sequence 5309, Application US/09783590
; Patent No. US2002011085OA1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseline, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5309
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (42)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (64)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (98)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (134)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (139)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (232)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (257)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (263)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (270)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (272)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (312)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (319)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (326)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-5309

Query Match   3.8%;   Score 33; DB 9; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;
Matches 33; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   441 TCAACGAGGTGCTGAAATCTTGGCTGATACTTT 473

```
Db    65  |||||||||||||||||||||||||||||||||
          TCAACGAGGTGCTGAAATCTTGGCTGATACTTT  97
```

RESULT 10
US-10-246-785-16
; Sequence 16, Application US/10246785
; Publication No. US2003014844A1
; GENERAL INFORMATION:
; APPLICANT: IRM, LLC
; APPLICANT: The Scripps Research Institute
; APPLICANT: Liao, Jiayu
; APPLICANT: Sheng, Ding
; APPLICANT: Schultz, Peter G
; TITLE OF INVENTION: Sweet Taste Receptors
; FILE REFERENCE: 36-002810US/PC
; CURRENT APPLICATION NUMBER: US/10/246,785
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/323,450
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-246-785-16

Query Match   2.7%;   Score 23; DB 15; Length 2523;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   807  TGGGCTGCTCACCTGCTGCTCAG  829
          |||||||||||||||||||||||
Db    10  TGGGCTGCTCACCTGCTCTCAG   32
```

RESULT 11
US-09-361-652-4
; Sequence 4, Application US/09361652
; Publication No. US2003003663A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; APPLICANT: Ryba, Nick
; APPLICANT: Hoon, Mark
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; TITLE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-088610US
; CURRENT APPLICATION NUMBER: US/09/361,652
; CURRENT FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/094,465
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2771
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat G-protein coupled receptor B3 (GPCR-B3)
US-09-361-652-4

Query Match   2.7%;   Score 23; DB 10; Length 2771;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   807  TGGGCTGCTCACCTGCTGCTCAG  829
          |||||||||||||||||||||||
Db    66  TGGGCTGCTCACCTGCTGCTCAG  88
```

RESULT 12
US-09-927-315-4
; Sequence 4, Application US/09927315
; Publication No. US2003004004A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.
; APPLICANT: Chandrashekar, Jayaram
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-120110US
; CURRENT APPLICATION NUMBER: US/09/927,315
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2771
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat T1R1 sweet taste receptor
US-09-927-315-4

Query Match   2.7%;   Score 23; DB 10; Length 2771;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   807  TGGGCTGCTCACCTGCTGCTCAG  829
          |||||||||||||||||||||||
Db    66  TGGGCTGCTCACCTGCTGCTCAG  88
```

RESULT 13
US-10-190-417-4
; Sequence 4, Application US/10190417
; Publication No. US2003016613A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Chandrashekar, Jayaram
; APPLICANT: Hoon, Mark A.
; APPLICANT: Nelson, Greg
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet and Amino Acid Heterodimeric Taste
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: 02307E-120130US
; CURRENT APPLICATION NUMBER: US/10/190,417
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 09/927,315
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/358,925
; PRIOR FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2771
; TYPE: DNA
; ORGANISM: Rattus norvegicus

```
;   FEATURE:
;   OTHER INFORMATION: rat T1R1
US-10-190-417-4

Query Match          2.7%;   Score 23;  DB 15;  Length 2771;
Best Local Similarity  100.0%;  Pred. No. 0.19;
Matches     23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     807  TGGGCTGCTCACCTGCTGCTCAG 829
            |||||||||||||||||||||||
Db      66  TGGGCTGCTCACCTGCTGCTCAG 88


RESULT 14
US-09-764-847-1157
;   Sequence 1157, Application US/09764847
;   Patent No. US2002013276741
;   GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PC009
;   CURRENT APPLICATION NUMBER: US/09/764,847
;   CURRENT FILING DATE:  2001-01-17
;   Prior application data removed - consult PALM or file wrapper
;   NUMBER OF SEQ ID NOS: 2003
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 1157
;   LENGTH: 21470
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-847-1157

Query Match          2.3%;   Score 20;  DB 9;  Length 21470;
Best Local Similarity  100.0%;  Pred. No. 7.9;
Matches     20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     167  GCAAGCTGGGGCTGCAGGTG 186
            ||||||||||||||||||||
Db     509  GCAAGCTGGGGCTGCAGGTG 528


RESULT 15
US-10-092-154-1157
;   Sequence 1157, Application US/10092154
;   Publication No. US2003005437541
;   GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PC009C1
;   CURRENT APPLICATION NUMBER: US/10/092,154
;   CURRENT FILING DATE:  2002-03-07
;   NUMBER OF SEQ ID NOS: 2003
;   Prior Application removed - See File Wrapper or Palm
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 1157
;   LENGTH: 21470
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-092-154-1157

Query Match          2.3%;   Score 20;  DB 15;  Length 21470;
Best Local Similarity  100.0%;  Pred. No. 7.9;
Matches     20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     167  GCAAGCTGGGGCTGCAGGTG 186
            ||||||||||||||||||||
Db     509  GCAAGCTGGGGCTGCAGGTG 528


Search completed: June   7, 2004, 23:21:00
Job time : 429 secs
```

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2004, 21:31:51 ; Search time 378 Seconds
(without alignments)
3203.006 Million cell updates/sec

Title: US-09-889-733B-2
Perfect score: 1454
Sequence: 1 MAGLGHPSAFGRATHAVVRA.........EKVDGLLTCCSVLINKKVDS 285

Scoring table: BLOSUM62
                Xgapop 10.0 , Xgapext 0.5
                Ygapop 10.0 , Ygapext 0.5
                Fgapop 6.0 , Fgapext 7.0
                Delop   6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09889733/runat_07062004_083527_7960/app_query.fasta_1.455
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09889733 @CGN_1_1708 @runat_07062004_083527_7960 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Database :   N_Geneseq_29Jan04:*
             1: geneseqn19808:*
             2: geneseqn19908:*
             3: geneseqn20008:*
             4: geneseqn2001as:*
             5: geneseqn2001bs:*
             6: geneseqn2002s:*
             7: geneseqn2002as:*
             8: geneseqn2003bs:*
             9: geneseqn2002ozs:*
            10: geneseqn2004s:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1454	100.0	858	3	AAA47655	Aaa47655 Dimethyla
2	1439	99.0	858	9	ADE71465	Ade71465 Human Thr
3	1439	99.0	1633	6	AAH02907	Aah02907 Human she
4	1439	99.0	1633	6	ABQ88102	Abq88102 Human ost
5	1439	99.0	3932	6	ABZ35939	Abz35939 Human sec
6	1439	99.0	4208	6	AAL41494	Aal41494 Drug meta
7	1433	98.6	858	9	ADE71463	Ade71463 Human Met
8	1379	94.8	3008	9	ADB53014	Adb53014 Primary r
9	1379	94.8	3008	9	ADB79843	Adb79843 Rat N-G,N
10	1200	82.5	3778	9	ADE53981	Ade53981 Human pro
11	1064	73.2	3651	3	AAF16077	Aaf16077 Human pro
12	694	47.7	1228	3	AAD00677	Aad00677 Human Hyd
13	694	47.3	1376	2	AAZ52967	Aaz52967 Human pro
14	688	47.3	858	3	AAA47656	Aaa47656 Dimethyla
15	669.5	46.0	845	3	AAA99390	Aaa99390 Human mel
16	551.5	37.9	770	6	ABS76755	Abs76755 Frog embr
17	482	33.1	873	4	ABL03335	Abl03335 Drosophil
18	427.5	29.4	444	6	ABN89945	Abn89945 Mouse clo
19	423	29.1	3223	4	AAK73698	Aak73698 Human imm
20	423	29.1	28772	8	ABK83555	Abk83555 Human cDN
21	386	26.5	472	8	ACH21905	Ach21905 Human adu
22	382	26.3	452	8	ACH27659	Ach27659 Human adu
23	373.5	25.7	3643	9	ABL03334	Abl03334 Drosophil
24	355.5	24.2	655	5	ADD34806	Add34806 Mouse mit
25	342	23.5	403	8	ACH32042	Ach32042 Human end
26	330	22.7	862	6	ABQ47886	Abq47886 Oligonucl
27	330	22.7	862	6	ABQ47887	Abq47887 Oligonucl
28	330	22.7	898	6	ABQ41321	Abq41321 Oligonucl
29	330	22.7	898	6	ABQ41320	Abq41320 Oligonucl
30	314	21.6	862	6	ABQ47888	Abq47888 Oligonucl
31	314	21.6	862	6	ABQ47889	Abq47889 Oligonucl
32	314	21.6	898	6	ABQ41319	Abq41319 Oligonucl
33	314	21.6	898	6	ABQ41318	Abq41318 Oligonucl
34	301	20.7	765	3	AAA47658	Aaa47658 Dimethyla
35	286	19.7	777	3	AAA47657	Aaa47657 Dimethyla
36	265.5	18.3	83946	6	ABQ88101	Abq88101 Human ost
37	245.5	16.9	545	5	ADD34805	Add34805 Mouse mit
38	237.5	16.3	504	9	ADB57724	Adb57724 Toxicity-
39	239	16.0	239	6	AAF68469	Aaf68469 Human lun
40	233	16.0	239	6	ABK38380	Abk38380 cDNA enco
41	233	16.0	239	6	ACA10709	Aca10709 Human lun
42	233	16.0	239	7	ABX99660	Abx99660 Lung canc
43	233	16.0	239	10	ADE71675	Ade71675 Human lun
44	204	14.0	6373	5	ABA14642	Aba14642 Human ner
45	204	14.0	6381	5	ABA14641	Aba14641 Human ner

ALIGNMENTS

RESULT 1
AAA47655
ID   AAA47655 standard; cDNA; 858 BP.
XX
AC   AAA47655;
XX
DT   08-NOV-2000 (first entry)
XX
DE   Dimethylarginine dimethylaminohydrolase (DDAH) coding sequence.
XX
KW   Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH2;
KW   arginine deaminase; hyperlipidemia; renal failure; hypertension;
KW   restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;
KW   ischemia reperfusion injury; septic shock; multi organ failure;
KW   arthritis; skin disorders; inflammatory cardiac disease; migraine;
KW   infection; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..858
FT   /*tag= a
FT   /product= "DDAH1"
XX
PN   WO200044888-A2.
XX
PD   03-AUG-2000.
XX
PF   26-JAN-2000; 2000WO-GB000226.
XX
PR   26-JAN-1999; 99GB-00001705.
PR   04-JUN-1999; 99GB-00013066.

```
XX
PA   (UNLO ) UNIV COLLEGE LONDON.
XX
PI   Vallance PJT, Leiper JM, Whitley GSJ, Charles IG;
XX
DR   WPI; 2000-543392/49.
DR   P-PSDB; AAB01477.
XX
PT   Novel methylarginase polypeptides and polynucleotides, used to identify
PT   modulators of them, which are used in the treatment of e.g. cancer,
PT   hypertension, and bacterial infections.
XX
PS   Claim 1; Page 55-56; 68pp; English.
XX
CC   Nucleotides encoding methylarginase polypeptides, vectors comprising
CC   these nucleotides and the polypeptides themselves can be used in
CC   medicaments for the treatment of hyperlipidemia, renal failure,
CC   hypertension, restenosis after angioplasty, atherosclerosis,
CC   complications of heart failure, schizophrenia, multiple sclerosis or
CC   cancer. Modulators of the enzyme can be used in medicaments for the
CC   treatment of ischemia-reperfusion injury of the brain or heart, cancer,
CC   lethal hypertension in severe inflammatory conditions such as septic
CC   shock or multi-organ failure, or local and systemic inflammatory
CC   disorders including arthritis, skin disorders, inflammatory cardiac
CC   disease, migraine, or microbial or bacterial infection
XX
SQ   Sequence 858 BP; 217 A; 214 C; 255 G; 171 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score:                  2.52e-155          Length:          858
Percent Similarity:     100.00%            Matches:         285
Best Local Similarity:  100.00%            Conservative:    0
Query Match:                               Mismatches:      0
DB:                                        Indels:          0
                                           Gaps:            3

US-09-889-733B-2 (1-285) x AAA47655 (1-858)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaAlaValValArgAla 20
Db   1 ATGGCCGGCCTCGGCCACCCTCCGCCTTCGGCCGGCGCCACCCACGCCGTGGTGCGGGGCG 60

Qy  21 LeuProGluSerLeuCysGlnHisAlaAlaLeuArgSerAlaLysGlyGluValAspVal 40
Db  61 CTACCCGAGTCGCTCTGCCAGCACGCGGCTCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 120

Qy  41 AlaArgAlaGluArgGlnHisGlyLeuTyrValGlyValSerLysGlyLeuGlyLeu 60
Db 121 GCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGTCGGGCAGCAAGCTGGGGCTG 180

Qy  61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysAspPheValGluAspVal 80
Db 181 CAGGTGGTGGAGCTGCCCGGCCGACGAGAGCCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 240

Qy  81 AlaAlaValCysGluGluThrArgProGlyAlaProSerArgArgGlyLys 100
Db 241 GCCGTGGTGTGCGAGGAGACGCGGCCCTCATCACCCGACCCGGGCGCCGAGCCGGAGGAAG 300

Qy 101 GluValAspMetMetLysGluLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
Db 301 GAGGTTGACATGATGAAGGAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360

Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
Db 361 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420

Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGlyAlaGluIleLeuAlaAspThrPheLysLysAsp 160
Db 421 GGCCTTTCCAAAAGGACAAATCAACGACGTGCTGAAATCTTGGCTGATACTTTTAGGGAC 480

Qy 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 481 TATGCAGTGTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540

Qy 181 AlaGlyLyrProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
Db 541 GCTGGGCCTCACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAGATC 600

Qy 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspIleAlaAla 220
Db 601 ATGCAACAGATGAGTGACCACCGTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660

Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAA 720

Qy 241 GluTyrProGluSerAlaLysValTyrGluLyuLeuLysAspHisMetLeuIleProVal 260
Db 721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG 780

Qy 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
Db 781 AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC 840

Qy 281 LysLysValAspSer 285
Db 841 AAGAAGGTAGACTCC 855

RESULT 2
ADE71465
ID   ADE71465 standard; DNA; 858 BP.
XX
AC   ADE71465;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human Thr87 wild-type dimethylarginine dimethylaminohydrolase-1 DNA.
XX
KW   Human; dimethylarginine dimethylaminohydrolase-1; DDAH 1; chromosome 1;
KW   single nucleotide polymorphism; SNP; ADMA;
KW   asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
KW   coronary heart disease; cerebrovascular disease; hypertension; diabetes;
KW   susceptibility; genotyping; therapeutic targeting; antidiabetic;
KW   cardiant; cereboprotective; hypotensive; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..858
FT                   /tag= a
FT                   /product= "Wild-type (Thr87) DDAH 1"
FT                   replace(260,T)
FT                   /tag= b
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT                   /note= "Results in Thr (C variant) or Met (T variant) at
FT                   position 87 of the protein"
XX
PN   WO203089638-A1.
XX
PD   30-OCT-2003.
XX
PF   11-APR-2003; 2003WO-FI000274.
XX
PR   19-APR-2002; 2002US-00125456.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Valkonen V, Salonen JT, Pirskanen M, Tuomainen T, Laakso J;
PI   Laaksonen R;
XX
DR   WPI; 2003-854121/79.
DR   P-PSDB; ADE71466.
XX
PT   New nucleic acid encoding a variant dimethylarginine
PT   dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
PT   and its vascular complications, e.g. coronary or cerebrovascular disease
```

```
Alignment Scores:
Pred. No.:
Score:                1439.00          Length:        858-153
Percent Similarity:   99.65%           Matches:       283
Best Local Similarity: 99.30%          Conservative:  1
Query Match:          98.97%           Mismatches:    1
DB:                                    Indels:        0
                                       Gaps:          0

US-09-889-733B-2 (1-285) x ADE71465 (1-858)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db   1 ATGGGCGGGCTTGGGCATCCCAGCGCCTTCGGCCCGGCACCCAGCGGTGCGGGCG 60
Qy  21 LeuProGluSerLeuCysGlnHisAlaAlaLeuArgSerAlaLysSerAlaLysGluValAspVal 40
Db  61 CTACCCGAGTCGCTGTGCCAGCACGCGGCTGAGAAGCGCCAAGGGCGGAGGAGGTGGACGTC 120
Qy  41 AlaArgAlaGluArgGlnHisGlnHisGlnLeuTyrArgValLeuValValGlySerLeuGlyLeu 60
Db 121 GCCCGCGCGGAACGGCAGCAACGGCACCAGCTCTACGTGGGCGTGCTGGCAAGCTGGGGCTG 180
Qy  61 GlnValValGluLeuProAlaLeuAspAlaAspGluSerLeuProSerLeuProAspCysValPheValGluAspVal 80
Db 181 CAGGTGGTGGAGCTGCCGGCCGACGCAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 240
Qy  81 AlaValValCysGluGluThrAlaIleLeuThrAlaLeueuIleThrArgProGlyAlaProSerArgArgLys 100
Db 241 GCCGTGGTGTGCGAGGAGACGGCCCCCTTCATCACCCGACCCCGGGCGCCGAGCCGGAGGAG 300
Qy 101 GluValAspValGluAlaLeuAlaLeuGluGlyLysLeuGlnLeuGluAsnIleValGlyMetLys 120
```

```
Db 301 GAGGTTGACATGATGAAAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
Qy 121 AspGluAsnAlaThrLeuAspPheGlyGlyLysAspValLeuPheThrGlyArgGluPheVal 140
Db 361 GATGAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 420
Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyLysArgGlyValAlaGluIleLeuAlaAspThrPheLysAsp 160
Db 421 GGCCTTTCCAAAGACAAATCAACGAGGTGCTGAAATCTTGGCTATACTTTTAAGGAC 480
Qy 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540
Qy 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLeuAlaLeuLysIle 200
Db 541 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 600
Qy 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 601 ATGCAACAGATGAGTGACCACCGTACGACAAACTCACTGTGCCGTGATGACATAGCAGCA 660
Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValHisLeuLeuHisArgThrProGlu 240
Db 661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGGTTCTGCTGCACCGAACCCCGGAA 720
Qy 241 GluTyrProGluSerAlaLysValTyrGlyLysLeuGlyLysAspHisMetLeuIleProVal 260
Db 721 GAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCCGTG 780
Qy 261 SerMetSerGluLeuGlyLysValAlaAspGlyLeuLeuThrCysSerSerValLeuIleAsn 280
Db 781 AGCATGTCTGAACTGGGAAAAGGTTGGATGGGCTGCTCACCTGCTTGCTCAGTTTTAATTAAC 840
Qy 281 LysLysValAspSer 285
Db 841 AAGAAAGTAGACTCC 855
```

SQ Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:
Score: 3.22e-153   Length: 1633
                      1439.00   Matches: 283
Percent Similarity: 99.65%   Conservative: 1
Best Local Similarity: 99.30%   Mismatches: 1
Query Match: 98.97%   Indels: 0
DB: Gaps: 0

US-09-889-733B-2 (1-285) x AAH02907 (1-1633)

```
Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla  20
         ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db 323 ATGGCCGGGCTCGGCCACCCGCCTTCGGCCCGGGCCACCCACCGCCGGTGCGGGGCG    382

Qy  21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAspVal     40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 383 CTACCCGAGTCGCTCGGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC  442

Qy  41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 443 GCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG  502

Qy  61 GlnValValGlyLeuProAlaSerGluProAspCysValPheValGluAspAspVal     80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 503 CAGGTGGTGGAGCTGCCCGGCCCGACGCGAGCGCCTTCCGGACTCCGTCTTCGTGGAGGACGTG 562

Qy  81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys  100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 563 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAAG  622

Qy 101 GluValAspMetMetLysAspGluAlaLeuGluLysLeuAsnIleValGluMetLys     120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 623 GAGGTTGACATGATGAAAGATGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 682

Qy 121 AspGluAsnAlaThrLeuAspGlyGlyValGlyAspValLeuPheThrGlyArgGluPheVal 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 683 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTG  742

Qy 141 GlyLeuSerLysThrAsnGlnArgGlyAlaGluIleLeuAlaGluIleAspThrPheLysAspAsp 160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 743 GGCCTTTCCAAAAGCACAAATCAACGAGGTGCTGAAATCTTGGCTGAGAACTTTTAAGGAC  802

Qy 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysLeuSerLysSerPheCysSerMet 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 803 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG  862

Qy 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle  200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 863 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAAATCTGCACAGAGAGCCCTTAAGATC  922

Qy 201 MetGlnMetSerAspHisArgThrTyrAspLysLeuThrValProAspAspIleAlaAla  220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 923 ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA  982

Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisValArgThrProGlu 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 983 AACTGTATATATCTAAATATCCCACAAAGGGCACGTTCTTGCTGCACGAACCCCGGAA  1042

Qy 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysLeuAspHisMetLeuIleProVal 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1043 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGAAACTGAAGGACCATATGCTGATCCCCGTG 1102
```

```
Qy 261 SerMetSerGluLeuGlyLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn  280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1103 AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC 1162

Qy 281 LysLysValAspSer  285
         |||||||||||||||
Db 1163 AAGAAAGTAGACTCC  1177
```

Alignment Scores:
Pred. No.:
Score: 3.22e-153   Length: 1633
                      1439.00   Matches: 283
Percent Similarity: 99.65%   Conservative: 1
Best Local Similarity: 99.30%   Mismatches: 1
Query Match: 98.97%   Indels: 0
DB: Gaps: 0

KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
KW anti-inflammatory; immunosuppressive; neuroprotective; nootropic;
KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
KW secretory polynucleotide; secretory protein; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200283876-A2.
XX
PD 24-OCT-2002.
XX
PF 27-MAR-2002; 2002WO-US009921.
XX
PR 29-MAR-2001; 2001US-0280067P.
PR 29-MAR-2001; 2001US-0280068P.
PR 16-MAY-2001; 2001US-0291280P.
PR 17-MAY-2001; 2001US-0291829P.
PR 17-MAY-2001; 2001US-0291849P.
PR 19-JUN-2001; 2001US-0299428P.
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300001P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
DR WPI; 2003-075543/07.
DR P-PSDB; ABP75492.
XX
PT New human secretory proteins and polynucleotides, useful for diagnosing,
PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
PT neurological disorders (e.g. Alzheimer's), or cell proliferations or
PT cancers.
XX
PS Claim 1; SEQ ID NO 103; 458pp + Sequence Listing; English.
XX
CC The invention relates to a secretory polynucleotide (designated sptm)
CC comprising any of 567 polynucleotide sequences (ABZ35937-ABZ36403), a
CC naturally occurring polynucleotide sequence at least 90 % identical to
CC the polynucleotide sequence, a polynucleotide complementary to them or an
CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
CC treating, preventing or diagnosing a disease or condition associated with
CC the expression of functional SPTM. These are particularly useful for
CC diagnosing, treating or preventing autoimmune/inflammatory disorders
CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
CC schizophrenia or amnesia, or cell proliferative disorders (e.g.
CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
CC breast, cervix or prostate). Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3932 BP; 1089 A; 883 C; 775 G; 1185 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score: 1.13e-152   Length: 3932
Percent Similarity: 1439.00   Matches: 283
Best Local Similarity: 99.65%   Conservative: 1
Query Match: 99.30%   Mismatches: 1
DB: 98.97%   Indels: 0
                            7   Gaps: 0

US-09-889-733B-2 (1-285) x ABQ88102 (1-1633)

```
Qy    1  MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValArgAla    20
            |||||||||||||||||:||:||||||||||||||||||||||||||
Db  323  ATGGCCGGGCTCGGCCACCCGAGCGCCTTCGGCCGGGCCACCCACGCCGTGGTGCGGGCG   382

Qy   21  LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLeuGlyGluGluValAspVal    40
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  383  CTACCCGAGTCGCTGTGCCAGCACGCGCTGAGGAGCGCCCAAGGGCGAGGAGGTGGACGTC   442

Qy   41  AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu    60
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  443  GCCCGCGCGGGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG   502

Qy   61  GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal    80
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  503  CAGGTGGTGGAGCTGCCCGCGGACGAGAGCCCTTCCGGACGTGCGTCTTCGTGGAGGACGTG   562

Qy   81  AlaValValCysGluGluThrArgAlaLeuIleThrArgProGlyAlaProSerArgArgLys   100
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  563  GCCGTGGTGTGCGAGGAGACGCGGGCCCTCATCACCCGACCCGGCGCCGCCAGCAGGAAG   622

Qy  101  GluValAspMetMetLysGlyGluAlaLeuGluLysGluGlnLeuAsnIleValGluMetLys   120
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  623  GAGGTTGACATGATGAAAGGAGAAGCATTAGAAAAAGAACTTCAGCTCAATATAGTAGAAGATGAAA   682

Qy  121  AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal   140
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  683  GATGAAAATGCAACTTTAGATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTG   742

Qy  141  GlyLeuSerLysArgThrAsnGlnArgGlyAlaGlyIleLeuGluLeuGluLeuAspPheThrLeuLysAsp   160
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  743  GGGCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC   802

Qy  161  TyrAlaValSerThrArgValProValAlaAspGlyValLeuHisLeuLysSerPheCysSerSerMet   180
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  803  TATGCAGTCTCCACAGTGCCAGTGGCCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG   862

Qy  181  AlaGlyProAsnLeuLeuAlaIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle   200
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  863  GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC   922

Qy  201  MetGlnGlnMetSerArgHisArgTyrArgAspTyrLysLeuThrValProAspAspAsnIleAlaAla   220
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  923  ATGCAACAGATGAGTGACCACCGTACGACAACAACTCACTGTGCCTGATGACATAGCAGCA   982

Qy  221  AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu   240
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  983  AACTGTATATATCTAAATATCCCAACAAAGGGACACGTGTTCTGCACCGAACCCCCGGAA  1042

Qy  241  GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal   260
            |||||||||||||||||||||||||||||||||||||||||||||||
Db 1043  GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTTAAGGACCATATGCTGATCCCCGTG  1102

Qy  261  SerMetSerGluLeuGluLysValLysLeuValAspGlyLeuLeuThrCysCysSerValLeuIleLeuAsn   280
            |||||||||||||||||||||||||||||||||||||||||||||||
Db 1103  AGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC  1162

Qy  281  LysLysValAspSer   285
            |||||||||||||||
Db 1163  AAGAAAGTAGACTCC  1177
```

RESULT 5
ABZ35939
ID   ABZ35939 standard; cDNA; 3932 BP.
XX
AC   ABZ35939;
XX
DT   10-FEB-2003 (first entry)
XX
DE   Human secretory polynucleotide SPTM SEQ ID NO 103.
XX

US-09-889-733B-2 (1-285) x ABZ35939 (1-3932)

```
Qy    1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ATGGCCGGGCTCGGGCACCCCTCGGCCCCGGCCACCCACGCCGTGGTGCGGGCG 76

Qy   21 LeuProGluSerLeuCysGlnHisHisAlaAlaLeuIleLeuArgSerAlaLysGlyGluValAspVal 40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 CTACCCGAGTCGCTCGGCCAGCACGGCGCTGAGAAGGCCAGAGGGCGAGGAGGTGGACGTC 136

Qy   41 AlaArgAlaGluArgAlaArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysGlyLeu 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 GCCCGCGCGGAACGGGCAGCACGGCACCAGCTCTACGTGGGCGTGCTGGGCAGCTCGGGCTG 196

Qy   61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  197 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 256

Qy   81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  257 GCCGTGGTGTGTGAGGAGACGAGGGCCCTCATCACCCGGCCCGGGGCCCCAGCCCGGAGGAAG 316

Qy  101 GluValAspMetMetLysGlyLeuGlnLeuAsnIleValGluMetLys 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  317 GAGGTTGACATGATGAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 376

Qy  121 AspGluAsnAlaThrLeuAspGlyGlyTyrAspValLeuPheThrHisGlyArgGluPheVal 140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  377 GATGAAAATGCAACTTAGATGGCCGGAGATGTTTATTCACAGAGCAGAGAATTTTTTGTG 436

Qy  141 GlyLeuSerLysArgThrAsnGlnArgGlyHisAlaGlyIleLeuAlaAlaAspThrPheLysAsp 160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  437 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 496

Qy  161 TyrAlaValSerThrValProValAlaAlaAspHisGlyLeuHisLeuLysSerPheCysSerSerMet 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  497 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 556

Qy  181 AlaGlyProAsnLeuIleAlaIleLeuGlySerSerGluSerArgAlaGlnLysAlaLeuIle 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  557 GCTGGGCCTACCTGATCGCAATTGGGTCTAGTGAGAATCTGCACAGAAGGCCCTTAAGATC 616

Qy  201 MetGlnGlnMetSerAspHisArgTyrArgAspLysLeuThrValProAspAspIleAlaAla 220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  617 ATGCAACAGATGAGTGACACCACCGCTACGACAAACTCACTGTGCCCTGATGACATAGCAGCA 676

Qy  221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  677 AACTGTATATATCTAAATATCCCCACCAAAGGGCACGTCTTGCTGCACCGAACCCCGGAA 736

Qy  241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuProVal 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  737 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG 796

Qy  261 SerMetSerGluLeuGluLysValAspProGlyLeuLeuThrCysCysSerValLeuIleAsn 280
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  797 AGCATGTCTGAACTGGAAAAGGTGGATCCGGGCTGCTCACCTGCTGCTCAGTTTAATTAAC 856

Qy  281 LysLysValAspSer 285
        |||||||||||||||
Db  857 AAGAAAGTAGACTCC 871
```

RESULT 6
AAL41494
ID   AAL41494 standard; DNA; 4208 BP.
XX
AC   AAL41494;
XX
DT   21-NOV-2002 (first entry)
XX
DE   Drug metabolising enzyme encoding DNA - 7671089CB1.
XX
KW   Anti-HIV; antiarteriosclerotic; dermatological; cytostatic; thyromimetic;

```
CC   metabolising enzyme of the invention
XX
SQ   Sequence 4208 BP; 1092 A; 1017 C; 883 G; 1216 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                                    4208
Score:                 1.25e-152
Score:                 1439.00                Matches:        283
Percent Similarity:    99.65%                 Conservative:   1
Best Local Similarity: 99.30%                 Mismatches:     1
Query Match:           98.97%                 Indels:         0
DB:                                           Gaps:           0

US-09-889-733b-2 (1-285) x AAL41494 (1-4208)

Qy     1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaAlaValArgAla 20
           ||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db   368 ATGGCCGGGCTCGGCCACCCGCGCCTTCGGCCGGGCCACCCACGCCGTGGTGCGGGCG 427

Qy    21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAspVal 40
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 CTACCCGAGTCGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCCAGGAGGTGGACGTC 487

Qy    41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 GCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 547

Qy    61 GlnValValGluLeuProAlaAspGluSerLeuProCysValPheValGluAspVal 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 CAGGTGGTGGAGCTGCCCGGACGAGGAGCCCTTCCGGACTCGGTCTTCGTGGAGGACGTG 607

Qy    81 AlaValValCysGluValThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAAG 667

Qy   101 GluValAspMetMetLysLysGluGluLeuGlnLeuAsnIleValGluMetLys 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 GAGGTTGACATGATGAAAGAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 727

Qy   121 AspGluAsnAlaThrLeuAspGlyGlyAspValAspValLeuPheThrGlyArgGluPhePheVal 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 GATGAAAATGCAACTTTAGATGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTG 787

Qy   141 GlyLeuSerLysArgThrAsnGlnArgGlyValAlaGluIleLeuAlaAspThrPheLysAsp 160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 GGCTTTCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 847

Qy   161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGTTGCATTTGAAGAGTTTTCTGCAGCATG 907

Qy   181 AlaGlyProAsnLeuIleLeuAlaIleGlySerSerAlaGlnGlnLysLysAlaLeuLysIle 200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GCTGGGCCTAACCTGATCCGATCGCAATTGGGTCTAGTGAAACTGCACAGAAGGCCCTTAAGATC 967

Qy   201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 ATGCAACAGATGAGTGACCACCGTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 1027

Qy   221 AsnCysIleTyrIleuAsnIleProAsnIleLeuValLeuLeuHisValAspThrProGlu 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028 AACTGTATATATATCTAAATATCCCAACAAAGGGCACGTTCTTGCTGCACCGGAACCCCGGAA 1087

Qy   241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuGluProVal 260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG 1147

Qy   261 SerMetSerGluLeuGluLysGluLeuThrCysCysSerValLeuIleAsn 280
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148 AGCATGTCTGAACTGGAAAAGTGGATGGGCGTGCTCACCTGCTGCAGTTTTAATTAAC 1207

Qy   281 LysLysValAspSer 285
           |||||||||||||||
Db  1208 AAGAAAGTAGACTCC 1222
```

RESULT 7
ADE71463
ID   ADE71463 standard; DNA; 858 BP.
XX
AC   ADE71463;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human Met87 variant dimethylarginine dimethylaminohydrolase-1 DNA.
XX
KW   Human; variant dimethylarginine dimethylaminohydrolase-1; variant DDAH 1;
KW   chromosome 1; single nucleotide polymorphism; SNP; ADMA;
KW   asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
KW   coronary heart disease; cerebrovascular disease; hypertension; diabetes;
KW   susceptibility; genotyping; therapeutic targeting; antidiabetic;
KW   cardiant; cerebroprotective; hypotensive; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..858
FT   /tag= a
FT   /product= "Met87 variant DDAH 1"
FT   variation   replace(260,C)
FT   /tag= b
FT   /standard_name= "single nucleotide polymorphism (SNP)"
FT   /notes "Results in Thr (C variant) or Met (T variant) at
FT   position 87 of the protein"
XX
PN   WO200309638-A1.
XX
PD   30-OCT-2003.
XX
PF   11-APR-2003; 2003WO-FI000274.
XX
PR   19-APR-2002; 2002US-00125456.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Valkonen V,   Salonen JT,   Pirskanen M,   Tuomainen T,   Laakso J;
PI   Laaksonen R;
XX
DR   WPI; 2003-854121/79.
DR   P-PSDB; ADE71464.
XX
PT   New nucleic acid encoding a variant dimethylarginine
PT   dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
PT   and its vascular complications, e.g. coronary or cerebrovascular disease
PT   or hypertension.
XX
PS   Claim 4; SEQ ID NO 1; 37pp; English.
XX
CC   The invention relates to a variant dimethylarginine
CC   dimethylaminohydrolase-1 (DDAH 1; ADE71464) and a nucleic acid encoding
CC   it (ADE71463). DDAHs regulate the metabolism of ADMA (asymmetrical
CC   dimethyl-arginine; a naturally occurring nitric oxide synthase inhibitor)
CC   and nitric oxide, which are relevant in disorders such as cardiovascular
CC   disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
CC   has been implicated in susceptibility to diseases such as familial
CC   combined hyperlipidaemia, premature coronary artery disease (CAD), non
CC   insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
CC   The variant DDAH 1 of the invention has Met at position 87, rather than
CC   Thr, which is caused by a C to T change in exon 1 at a position
CC   corresponding to base 260 in ADE71463 and ADE71465. Individuals hetero-
CC   or homozygous for the Thr87Met allele have an increased risk of
CC   developing cardiovascular diseases and diabetes. The invention also
CC   relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a
CC   susceptibility to a cardiovascular disease and diabetes in an individual
CC   by determining the DDAH 1 genotype; a method of treating diabetes or
CC   vascular complications of diabetes by enhancing nitric oxide
CC   availability, production or concenration; a method of targeting
CC   treatment of cardiovascular disease and diabetes in a hypertensive

```
CC  patient by determining the DDAH 1 genotype and treating them with a drug
CC  that affects nitric oxide availability, production or metabolism; and a
CC  transgenic animal which carries a human variant DDAH 1 nucleic acid
CC  sequence. The nucleic acid molecules and polypeptides are useful for
CC  treating cardiovascular disease (e.g., coronary heart disease;
CC  cerebrovascular disease, and hypertension), and diabetes and its vascular
CC  complications. The methods are useful for determining whether a patient
CC  will benefit from treatment with a drug which affects nitric oxide
CC  availability, production or metabolism; a drug which reduces ADMA
CC  availability or concentration; or an agent which elevates DDAH
CC  availability or concentration (such as DDAH agonist). The methods are
CC  also useful for determining whether a patient will be at risk of adverse
CC  effects if DDAH antagonists are administered. The present sequence
CC  represents a specifically claimed nucleic acid encoding the variant
CC  (Met87) DDAH 1.
XX
SQ  Sequence 858 BP; 218 A; 212 C; 258 G; 170 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:              6.14e-153          Length:          858
Score:                 1433.00            Matches:         282
Percent Similarity:    99.30%             Conservative:    1
Best Local Similarity: 98.95%             Mismatches:      2
Query Match:           98.56%             Indels:          0
DB:                                       Gaps:            0

US-09-889-733B-2 (1-285) x ADB71463 (1-858)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla    20
          |||||||||||||||||||||||||||||||||||||||||::::||
Db   1 ATGGCCGGCTCGGGCCACCCCTCGGCCCGCCTTCGCGGGGCCACCCACGCCGTGGTGCGGGCG 60

Qy  21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAspVal      40
Db  61 CTACCCAGTCGCTGCTCGGCCAGCACGCGGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 120

Qy  41 AlaArgAlaGluArgGlnHisGlnHisLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
Db 121 GCCCGCGCCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCTGGGCGTG      180

Qy  61 GlnValValGluLeuProAlaSerLeuProAspCysValPheValGlnValAspVal      80
Db 181 CAGGTGGTGGAGCTGCCCGGCCCAGCGAGCCCTTCCGGACTGCGTCTTCGTGGAGGACGTG  240

Qy  81 AlaAlaValCysGluGluThrArgAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
Db 241 GCCGTGGTGTGCGAGGAGATGGCCCCTCATCACCCGACCCGGGCCCCGGCAGCCGGAGGAAG  300

Qy 101 GluValAspMetMetLysGluAsnIleValLeuGlnLeuAsnIleValGluMetLys     120
Db 301 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATAGTAGAGATGAAA     360

Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspGlyAspMetGlyPheThrGlyArgGluPhePheVal 140
Db 361 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGACAGAGAATTTTTTGTG   420

Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuGluAlaAspThrPheLysLysAsp 160
Db 421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC    480

Qy 161 TyrAlaValSerThrValProValAlaAspGlyHisLeuHisLeuSerPheCysSerSerMet 180
Db 481 TATGCAGTCTCCACAGTGCCCAGTGCCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG   540

Qy 181 AlaGlyProAsnLeuIleAlaIleAlaGlySerGluSerGlySerAlaGlnLysAlaLeuLysIle 200
Db 541 GCTGGGCCCTAACCTGATCGCAATTGGTGGTTCTGAATCTGGAGGCCCTTAAGATC      600

Qy 201 MetGlnGlnMetSerAspHisArgTyrAspHisTyrAspHisLeuThrValProAspAspIleAlaAla 220
Db 601 ATGCAACAGATGAGTGACCACCGCTACGACCAAACTCACTGTGCCTGATGACATAGCAGCA   660

Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisValGlyArgThrProGlu 240
Db 661 AACTGTATATCTAAAATCCCACAACAAGGGCACAGTGCTTGCTGCACCGAACCCCGGAA    720

Qy 241 GluTyrProGluSerAlaLysValTyrGlyLeuLysLeuAspHisMetLeuIleProVal    260
Db 721 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG   780

Qy 261 SerMetSerGluLeuGlyLysValAspGlyLeuGluThrCysCysSerValLeuIleAsn    280
Db 781 AGCATGTCGAACTGGAAAGGTGGATGGGCTGCTCCACCTGCTGCTCAGTTTTAATTAAC    840

Qy 281 LysLysValAspSer 285
Db 841 AAGAAAGTAGACTCC 855
```

```
RESULT 8
ADB53014
ID  ADB53014 standard; DNA; 3008 BP.
XX
AC  ADB53014;
XX
DT  04-DEC-2003  (first entry)
XX
DE  Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3556.
XX
KW  toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW  toxicity marker; toxicity progression; drug screening;
KW  primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS  Rattus norvegicus.
XX
PN  WO2003065993-A2.
XX
PD  14-AUG-2003.
XX
PF  04-FEB-2003; 2003WO-US003482.
XX
PR  04-FEB-2002; 2002US-0353171P.
PR  13-MAR-2002; 2002US-0363534P.
PR  08-APR-2002; 2002US-0370248P.
PR  10-APR-2002; 2002US-0371134P.
PR  10-APR-2002; 2002US-0371135P.
PR  10-APR-2002; 2002US-0371150P.
PR  11-APR-2002; 2002US-0371413P.
PR  19-APR-2002; 2002US-0373601P.
PR  19-APR-2002; 2002US-0373602P.
PR  22-APR-2002; 2002US-0374139P.
PR  08-MAY-2002; 2002US-0378370P.
PR  09-MAY-2002; 2002US-0378652P.
PR  09-MAY-2002; 2002US-0378653P.
PR  09-MAY-2002; 2002US-0378665P.
PR  09-JUL-2002; 2002US-0394230P.
PR  09-JUL-2002; 2002US-0394253P.
PR  04-SEP-2002; 2002US-0407688P.
PR  28-JAN-2003; 2002US-0442900P.
XX
PA  (GENE-) GENE LOGIC INC.
XX
PI  Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI  Elashoff M;
XX
DR  WPI; 2003-731472/69.
XX
PT  Determining if a compound induces a toxic effect on a tissue or cell, for
PT  identifying hepatotoxic compounds, comprises comparing a gene expression
PT  profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT  mean values.
XX
PS  Claim 44; SEQ ID NO 3556; 874pp; English.
XX
CC  The present invention describes a method for determining whether a
CC  compound induces a toxic effect on a tissue or cell. The method comprises
CC  preparing a gene expression profile of a tissue or cell sample exposed to
```

CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:   Length: 3008
Score: 5.09e-146   Matches: 268
                         1379.00   Conservative: 8
Percent Similarity: 96.84%   Mismatches: 9
Best Local Similarity: 94.04%   Indels: 0
Query Match: 94.84%   Gaps: 0
DB:

US-09-889-733B-2 (1-285) x ADB53014 (1-3008)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyTyrArgAlaThrHisAlaValValArgAla 20
Db   432 ATGGCCGGCCTCAGCCACCCCTCGGCTTCGGCGGGGCCACCCACCGCGTGGTGCGGGCT 491

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyValGluValAspVal 40
Db   492 CCGCCCGAGTCCCTGTGCCGCCACGCGCTCCCAGGGGCGGAAGAGGTGGATTTC 551

Qy   41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysGlyBleuGlyLeu 60
Db   552 GCTCGCGCTGAGCGCCAGCACCAGCTTACGTGGGCGTGCTGGGCAGCAAGTGGGGCTG 611

Qy   61 GlnValValGlnLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db   612 CAGGTGGTGCAGCTGCCGGCCGACGAGAGCCCTGCCTGACGTCGTGTTCGTGGAGGACGTG 671

Qy   81 AlaValValCysGluGluThrAlaLeuIleLeuThrArgProGlyAlaProSerArgArgLys 100
Db   672 GCCGTGGTGTGCGAGGAGACCGCCCTCATCACCCGCCCGGGCGCTAGCCGCAGGAAG 731

Qy   101 GluValAspMetMetLysGluAlaLeuGluLysGlnLeuGlnLeuAsnIleValGluMetLys 120
Db   732 GAGGTTGACATGATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAA 791

Qy   121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
Db   792 GATGAAAATGCAACCTTAGATGGTGGGGACGTCCTATTCACAGGCAGAGAGTTTTTTGTG 851

Qy   141 GlyLeuSerLysArgThrAsnGlnAsnArgGlyAlaAlaIleLeuLeuAlaAspThrProLysLysAsp 160
Db   852 GGCCTTTCAAAAGACAAATCAACGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGAC 911

Qy   161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db   912 TACGGAGTTTCCACAGTGCCAGTTGCCGATGGCCTTCATTTAAAGAGTTTCTGCAGCATG 971

Qy   181 AlaGlyProAsnLeuIleLeuAlaIleGlySerSerGluSerAlaGlnLysLysAlaLeuLysIle 200
Db   972 GCTGGGCCCCAACCTGATCCTGGCAATAGGGTCCAGTGAATCTGCCCAGAAGGCCCTCAAGATC 1031

Qy   201 MetGlnIleMetSerAspHisArgTyrArgLysLeuThrValProAspAspAspIleAlaAla 220
Db 1032 ATGCAACAGATGAGTGACCACCGGTTATGACAAGCTCACTGTACCGGACGACATGGCCCCC 1091

Qy   221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 1092 AACTGTATATATTTAAATATCCCCAGCAAAGGGCATGTCTTGTCACCGAACCCCAGAA 1151

Qy   241 GluTyrProGluSerAlaLysValTyrGluLysLeuGluLysAspHisMetLeuIleProVal 260

Db 1152 GAGTACCCAGAGAAGCCCAAAGTTTATGAGAAGCTCAAGGACCATCTACTGATCCCTGTG 1211

Qy   261 SerMetSerGluLysGluLysGlyLeuValLeuThrCysCysValLeuLeuGluAsn 280
Db 1212 AGCAATTCTGAGATGCAGAAAAGGTGGACAGGCTTGCTCACCTGCTGTCCGTTTTATTAAC 1271

Qy   281 LysLysValAspSer 285
Db 1272 AAGAAGACAGACTCT 1286

RESULT 9
ADB79843
ID   ADB79843 standard; DNA; 3008 BP.
XX
AC   ADB79843;
XX
DT   04-DEC-2003 (first entry)
XX
DE   Rat N-G,N-G-dimethylarginine dimethylaminohydrolase DNA, SEQ ID 83.
XX
KW   Analgesic; pain; streptozocin-induced diabetes; rat; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   EP1279744-A2.
XX
PD   29-JAN-2003.
XX
PF   26-JUL-2002; 2002EP-00255249.
XX
PR   27-JUL-2001; 2001GB-00018354.
PR   07-FEB-2002; 2002GB-00002910.
XX
PA   (WARN) WARNER LAMBERT CO.
XX
PI   Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
XX
DR   WPI; 2003-395407/38.
DR   P-PSDB; ADB79842.
XX
PT   Use of isolated gene sequences and encoded polypeptides that are
PT   upregulated in the spinal cord in response to streptozocin-induced
PT   diabetes for screening compounds for the treatment of pain, or for
PT   diagnosing pain.
XX
PS   Claim 1; Page 155; 334pp; English.
XX
CC   The present invention relates to nucleotide sequences which are useful in
CC   the screening of compounds for the treatment of pain, or for the
CC   diagnosis of pain. The nucleotide sequences are up-regulated in the
CC   spinal cord in response to streptozocin-induced diabetes. The present
CC   sequence is one such nucleotide sequence.
XX
SQ   Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:   Length: 3008
Score: 5.09e-146   Matches: 268
                         1379.00   Conservative: 8
Percent Similarity: 96.84%   Mismatches: 9
Best Local Similarity: 94.04%   Indels: 9
Query Match: 94.84%   Gaps: 0
DB:

US-09-889-733B-2 (1-285) x ADB79843 (1-3008)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyTyrArgAlaThrHisAlaValValArgAla 20
Db   432 ATGGCCGGCCTCAGCCACCCCTCGGCTTCGGCGGGGCCACCCACCGCGTGGTGCGGGCT 491

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyValGluValAspVal 40
Db   492 CCGCCCGAGTCCCTGTGCCGCCACCGCGCTCCCAGGGGCGGAAGAGGTGGATTTC 551

XX
PR   31-MAY-2001; 2001US-0295048P.
XX
PA   (FARI/) FARIS M.
PA   (PEAR/) PEARSON C I.
XX
PI   Faris M, Pearson C I;
XX
DR   WPI; 2003-831619/77.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   prostate cancer, useful for diagnosing, treating or monitoring the
PT   progression of treatment of prostate cancer.
XX
PS   Claim 1; SEQ ID NO 328; 42pp; English.
XX
CC   The invention relates to a combination comprising a number of cDNAs
CC   expressed in prostate cancer. The invention also relates to a method for
CC   detecting differential expression of one or more cDNAs in a sample
CC   containing nucleic acids by hybridising a substrate with the nucleic
CC   acids, thus forming one or more hybridisation complexes, detecting
CC   hybridisation complex formation and comparing the complexes formed with
CC   standard complexes, where differences between the standard and the sample
CC   complex formation indicate differential expression of cDNAs in the
CC   sample. The differential expression is diagnostic of prostate cancer. The
CC   invention also relates to proteins and antibodies related to the cDNAs.
CC   The combination is useful for diagnosing, treating or monitoring the
CC   progression of treatment of prostate cancer. The antibodies are useful
CC   for detecting prostate cancer. This sequence represents a human prostate
CC   cancer cDNA of the invention.
XX
SQ   Sequence 3778 BP; 1064 A; 821 C; 719 G; 1150 T; 0 U; 24 Other;

Alignment Scores:
Pred. No.:   1.5e-125   Length: 3778
Score: 1200.00   Matches: 238
Percent Similarity: 99.58%   Conservative: 0
Best Local Similarity: 99.58%   Mismatches: 0
Query Match: 82.53%   Indels: 1
DB: 9   Gaps: 0

US-09-889-733B-2 (1-285) x ADE53981 (1-3778)

Qy   48 GlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuProAla 67
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2   CAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCCTGCAGGTGGTGGAGCTGCCGGCC 61

Qy   68 AspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGluThr 87
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 GACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACG 121

Qy   88 AlaLeuIleThrArgProGlyAlaProSerArgArgArgLysGluValAspMetMetLysGlu 107
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 GCCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAGGAGGTTGACATGATGAAAGAA 181

Qy   108 AlaLeuGluLysGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAsp 127
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 GCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAACTTGAT 241

Qy   128 GlyGlyAspValLeuPheThrArgGluPheValGlyLeuSerLysArgThrAsn 147
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 GGCGGAGATGTTTTATTCACAGGCAGAGATTTTTGTGGGCCTTTCCAAAAGGACAAAT 301

Qy   148 -GlnArgGlyAlaGlyAlaAspThrPheLysAspTyrAlaValSerThrValPr 167
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 CCAACGGCTGCTGAAATCTTGGTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCC 361

Qy   167 oValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAl 187
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 AGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGTGGCCTAACCTGATCGC 421

Qy   187 aIleGlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHi 207
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   41   AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   552 GCTCGCGCCGAGCGCCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 611

Qy   61   GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
           ||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||||
Db   612 CAGGTGGTGCAGCTGCCGGCCGACGAGAGCCTGCCTGACTGCGTGTTCGTGGAGGACGTG 671

Qy   81   AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   672 GCCGTGGTGTGCGAGGAGACCGCCCTCATCACCCGCCCGGGGGCGCCCAGCCGCCAGGAAG 731

Qy   101 GluValAspMetMetLysGluAlaLeuGluLysGlnLeuAsnIleValGluMetLys 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   732 GAGGTTGACATGATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAA 791

Qy   121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrArgGluPhePheVal 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   792 GATGAAAATGCAACCTTAGATGGTGGGACGTCCATTCACAGGCAGAGAGTTTTTTGTG 851

Qy   141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGlyIleAspThrPheLysAspAsp 160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   852 GGCCTTTCCAAAAGGACAAAATCAACAGGGCTGCTGAGATCTTGGCTGTATACTTTCAAGGAC 911

Qy   161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   912 TACGCAGTTTCCACAGTCCCGTGGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCATG 971

Qy   181 AlaGlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIle 200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   972 GCTGGCCCCAACCTGATCGCAATAGGGTCCAGTGAATCTCGCGAGAGGCCCTCAAGATC 1031

Qy   201 MetGlnMetSerAspHisArgTyrAspPlysLeuThrValProAspAspIleAlaAla 220
           |||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db   1032 ATGCAACAGATGAGTGACCACCGTTATGACAAGCTCACTGTACCGGACGACGGCCGCC 1091

Qy   221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
           ||:|||::||||::|||||||||||||||||||||||||||||||||||||||||||||||
Db   1092 AACTGTATATATTTAAATATCCCAGCAAAGGGCATGTCTTGCTCACCGAACCCAGAA 1151

Qy   241 GluTyrProGluSerAlaLysValTyrGluLysLysAspHisMetLeuIleProVal 260
           |||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db   1152 GAGTACCCAGAAAGCGCAAAGGTTTATGAGAAGCTCAAGGACCACTACTGATCCCTGTG 1211

Qy   261 SerMetSerGluLeuGlyLysValAspGlyGlyLeuLeuThrCysCysSerValLeuAsn 280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1212 AGCAATTCTGAGATGGAAAAGGTGGACGGCCTTGCTCACCTGCTCTTTTTTATTAAC 1271

Qy   281 LysLysValAspSer 285
           |||||||||||||||
Db   1272 AAGAAGGACAGACTCT 1286

RESULT 10
ADE53981
ID   ADE53981 standard; cDNA; 3778 BP.
XX
AC   ADE53981;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human prostate cancer cDNA #328.
XX
KW   Human; prostate cancer; ss; cDNA combination; differential expression;
KW   gene.
XX
OS   Homo sapiens.
XX
PN   US2003190640-A1.
XX
PD   09-OCT-2003.
XX
PF   29-MAY-2002; 2002US-00252157.

XX
SQ   Sequence 3651 BP; 1016 A; 801 C; 684 G; 1141 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.:   Length: 3651
Score: 4.24e-110   Matches: 212
Percent Similarity: 1064.00   Conservative: 0
Best Local Similarity: 99.53%   Mismatches: 1
Query Match: 99.53%   Indels: 1
                       73.18%   Gaps: 0
DB: 3

US-09-889-733B-2 (1-285) x AAF16077 (1-3651)

```
Qy     73 AspCysValPheValGluAspValAlaValValCysGluGlnThrAlaLeuIleThrArg  92
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4 GACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGGGAGGAGGCCCTCATCACCCGA  63

Qy     93 ProGlyAlaProSerArgArgLysGluGluValAspMetMetLysGluALeuGluLysLeu  112
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 CCCGGGGCGCCGAGCCGGAGGAGGGTTGACATGATGAAAGAAGCATTAGAAAAACTT  123

Qy    113 GlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeu  132
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 CAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTTTA  183

Qy    133 PheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGlu  152
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 TTCACAGGCAGAGAATTTTTTGTGGGCCTTTCCAAAAGGACAAAATCAACGAGGTGCTGAA  243

Qy    153 IleLeuAlaAspThrPheLeuAspTyrAlaValSerThrValProValAlaAspGlyLeu  172
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 ATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGGGTG  303

Qy    173 HisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleLeuAlaIleGlySerSerGlu  192
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 CATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAA  363

Qy    193 SerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisArgTyrAspLysLysLeu  212
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 TCTGCACAGAAGGCCCCTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTC  423

Qy    213 ThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnLysGlyHis  232
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 ACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCCAACAAAGGCAC  483

Qy    233 ValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLysLeu  252
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 GTCTTGCTGCACGAACCCCGAAGAGTATCCAGAAAGTGCAAAGGTTTATCGAGAAACTG  543

Qy    253 LysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLysLeuLeu  272
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 AAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAAGGTGGATGGGCTGCTC  603

Qy    273 ThrCysCysSerValLeuIleAsnLysLysValAspSer  285
          ||||||||||||||||||||||||||||||||||||||
Db    604 ACCTGCTGTCA-GTTTTAATTAACAGAAAGTAGACTCC  641
```

RESULT 12
AAD00677
ID   AAD00677 standard; cDNA; 1228 BP.
XX
AC   AAD00677;
XX
DT   08-SEP-2000 (first entry)
XX
DE   Human Hydrolase protein-2 (HYDRL-2) encoding cDNA.
XX
KW   Hydrolase; HYDRL; human; clone 949738; cytostatic; immunosuppressive;
KW   antiinflammatory; neuroprotective; cerebroprotective; anticonvulsant;
KW   nephrotropic; antibody; agonist; antagonist; diagnosis; treatment;
KW   prevention; cell proliferative disorder; cancer; inflammation; AIDS;
KW   Acquired Immune Deficiency Syndrome; autoimmune/inflammatory disorder;
KW   neurological disorder; epilepsy; stroke; medullary sponge kidney;

```
Db    422 AATTGGGTCTAGTGAATCTGCACAGAAGGGCCCTTAAGATCATGCAACAGATGAGTGACCA  481
Qy    207 sArgTyrAspLysThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIl  227
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 CCCCTACGACAAACTCACTGTGCCTGATGACATAGCAGCCAAACTGTATATATCCAAATAT  541
Qy    227 eProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLy  247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 CCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGAAGAGTATCCAGAAAGTGCAAA  601
Qy    247 sValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluGluLy  267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602 GGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAA  661
Qy    267 sValAspGlyLysLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer  285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 GGTGGATGGGCTGCTCACCTGCTGTCAGTTTTAATTAACAGAAAGTAGACTCC  716
```

RESULT 11
AAF16077
ID   AAF16077 standard; cDNA; 3651 BP.
XX
AC   AAF16077;
XX
DT   13-MAR-2001 (first entry)
XX
DE   Human prostate cancer antigen nucleotide sequence SEQ ID NO:512.
XX
KW   Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW   neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW   vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW   antibacterial; gene therapy; neural; immune; reproductive; renal;
KW   gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW   wound; infectious disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055174-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005988.
XX
PR   12-MAR-1999;   99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2000-587513/55.
DR   P-PSDB; AAB56874.
XX
PT   Prostate cancer associated gene sequences, referred to as prostate cancer
PT   antigens, useful for treatment, prevention, and diagnosis of disorders
PT   such as prostate cancer.
XX
PS   Claim 1; Page 976-977; 2338pp; English.
XX
CC   AAF15566 to AAF16505 encode the human prostate cancer associated
CC   proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC   The prostate cancer antigens can have neuroprotective, cytostatic,
CC   cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC   nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC   and can be used in gene therapy. The prostate cancer antigen
CC   polynucleotides may be used for detection of prostate cancer, chromosome
CC   identification, as chromosome markers, and for numerous other diagnostic
CC   or research purposes. The prostate cancer antigens may be used to treat
CC   disorders such as neural, immune, muscular, reproductive,
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC   AAB57303 represent sequences used in the exemplification of the present
CC   invention
```

```
Db   320  CCAGAGAGCTGCCGCTGCGGAAGGTGCGGGGGCCGCCTTCCCGGTCTGGATCTGCC      379
          ||||||||||                                   :::|||||||||||
Qy    42  ArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLysGlyGln     61
          :::|||:::|||  :::
Db   380  AAAGCTCAAGGAGGACACGGGGTGCTGGAGGTAAACTGAGGCAACGACTGGGGCTACAG   439
Qy    62  ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla  81
          :::||||||||||||||:::||||||
Db   440  CTGCTAGAAACTGCCACCTGAGAGTCATTGCCGCTGGGACCGCTGCTTGGCCACACGGCC  499
Qy    82  ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu  101
          |||:::                  :::||||||||||||:::|||
Db   500  GTGATCCAAGGGACACCGGCCCCTAATCACGCGGCGCCCCGAGCCCGGTCGTAGGCCCAG  559
Qy   102  ValAspMetMetLysGluValAlaLeuGluGlyLysLeuGluLeuAsnIleValGlnMetLysArgAsp  121
Db   560  GTCGATGGAGGTCCCCAAAACCCCTGCAAGACCCTGGGGCTTCCGAATTGTGAAATAGGAGAC         619
Qy   122  GluAsnAlaThrLeuAspGlyGlyValLeuPheThrGlyArgGluPheValGly  141
          |||||                                  |||||||||
Db   620  GAGAACACGGACGCTGGATGGCACTGACGTTCTCTTCACCGGCCGGGAGTTTTCGTAGGC  679
Qy   142  LeuSerLysArgThrAsnGlnArgGlyValAlaGluIleLeuAlaAspThrPheLysAspTyr  161
          |||:::                             :::|||||||       :::|||||||
Db   680  CTCTCCAAATGGACCAATACCGGAGGAGCTGAGCAGTCGTGGCGGACACGTTCCGGGACTTC  739
Qy   162  AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla  181
          |||||||:::                    :::||||||||||||:::
Db   740  GCCGTCTCCACTGTGCCAGTTCTCGGGTCCTCCGGGTCTCTGCGGCATGGGG          799
Qy   182  GlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMet  201
          |||||||                   :::       :::|||||||:::|||||||
Db   800  GGACCTCGGCACTGTTGTGCGACGGCAGGCAGCAGCGCCCCAAAAGGCGTGTCCGGGCAATG  859
Qy   202  GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsn  221
          |||||||                                  :::|||||||
Db   860  GCAGTGCTGACAGATCACCCATATGCCTCCCTGACCCCCCAGATGACCCAGCTGCTGAC  919
Qy   222  CysIleTyrLeu------AsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro  239
          |||:::                                     :::|||||||
Db   920  TGTCTTCCTTCGTCCTGGGTTCGCCCCCTTCCTCCTGCACCACCGGGAGGT  979
Qy   240  GluGluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetIlePro  259
          |||||                         :::|||||||:::|||
Db   980  GGGGATCTGCCCAACGACCCAGGAGCCCAGAGAGCTTCTGATGTCACCCTGGTACCT  1039
Qy   260  ValSerMetSerGluLeuGluLysValAsp---GlyLeuLeuThrCysSerSerValLeu  278
          |||:::                         :::|||||||
Db  1040  GTGTCCTGCTCAGAACTGGAGAAGGCCGGCGGCCGCCGGCTCAGCTCCCTGCTTGGTGCTC  1099
```

Alignment Scores:

Pred. No.:			
Score:	9.99e-69	Length:	1228
	694.00	Matches:	146
Percent Similarity:	68.93%	Conservative:	47
Best Local Similarity:	52.14%	Mismatches:	83
Query Match:	47.73%	Indels:	4
DB:		Gaps:	3

US-09-889-733B-2 (1-285) x AAD00677 (1-1228)

```
Qy    3  GlyLeuGlyHisPro---SerAlaPheGlyArgAlaThrHisAlaValValArgAlaLeu  21
         |||:|||||||                    |||||||||::        :::
Db  260  GGGATGGGACGCGCCGGGGAGGGGCGTTGGGCCCGGCTACGTCATGCCCTGCTCCCTGCTCC  319
Qy   22  ProGluSerLeuCysGlnHisAlaLeuAlaLeuArgSerAlaLysGlyAlaLysGlyGluValAspValAla  41
```

Qy   202 GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspIleAlaAlaAsn 221
                ::::||| ||| |||||||||||| ::::: |||| || ||||
Db   894 GCAGTGCTGACAGATCACCCATATGCTCCTCCCTGACCTCCCAGATGACGACCTGCTGAC 953

Qy   222 CysIleTyrLeu-----AsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro 239
                ||| || |: | ||| :::||||||||| |||||||||| || |||
Db   954 TGTCTCTTCTTCGTCCTGGTTGCCTGGTGCCCCCTTCCTCTGCACCGTGGAGGT 1013

Qy   240 GluGluTyrProGluSerAlaLysValTyrGluLeuLysAspHisMetLeuIlePro 259
                ::| || ||||||||| ::: ||| :::||||||||||||| ||||
Db   1014 GGGGATCTGCCCAAACGCCAGGAGGCACTGCAGGAAGCTCTGATGTCACCCTGGTACCT 1073

Qy   260 ValSerMetSerGluLeuGluLysValAsp---GlyLeuLeuThrCysCysSerValLeu 278
                ||||||| |||||||||| :::||| |||||||||| :::: ||||||
Db   1074 GTGTCCCTGCTCAGAACTGGAGAGGGCTGGGCCGGGCTCAGTCCCCTGCTTGGTGCTC 1133

RESULT 14
AAA47656
ID   AAA47656 standard; cDNA; 858 BP.
XX
AC   AAA47656;
XX
DT   08-NOV-2000 (first entry)
XX
DE   Dimethylarginine dimethylaminohydrolase (DDAH2) coding sequence.
XX
KW   Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH2;
KW   arginine deaminase; hyperlipidemia; renal failure; hypertension;
KW   restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;
KW   ischemia reperfusion injury; septic shock; multi organ failure;
KW   arthritis; skin disorders; inflammatory cardiac disease; migraine;
KW   infection; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..858
FT   /*tag= a
FT   /product= "DDAH2"
XX
PN   WO200044888-A2.
XX
PD   03-AUG-2000.
XX
PF   26-JAN-2000; 2000WO-GB000226.
XX
PR   26-JAN-1999;   99GB-00001705.
PR   04-JUN-1999;   99GB-00013066.
XX
PA   (UNLO) UNIV COLLEGE LONDON.
XX
PI   Vallance PJT, Leiper JM, Whitley GSJ, Charles IG;
XX
DR   WPI; 2000-543392/49.
DR   P-PSDB; AAB01478.
XX
PT   Novel methylarginase polypeptides and polynucleotides, used to identify
PT   modulators of them, which are used in the treatment of e.g. cancer,
PT   hypertension, and bacterial infections.
XX
PS   Claim 1; Page 57-58; 68pp; English.
XX
CC   Nucleotides encoding methylarginase polypeptides, vectors comprising
CC   these nucleotides and the polypeptides themselves can be used in
CC   medicaments for the treatment of hyperlipidemia, renal failure,
CC   hypertension, restenosis after angioplasty, atherosclerosis,
CC   complications of heart failure, multiple sclerosis or
CC   cancer. Modulators of the enzyme can be used in medicaments for the
CC   treatment of ischemia-reperfusion injury of the brain or heart, cancer,
CC   lethal hypertension in severe inflammatory conditions such as septic
CC   shock or multi-organ failure, or local and systemic inflammatory
CC   disorders including arthritis, skin disorders, inflammatory cardiac
CC   disease, migraine, or microbial or bacterial infection

XX
PA   (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI   Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR   WPI; 1999-621386/54.
DR   P-PSDB; AAY74144, AAY74145, AAY74146.
XX
PT   New human nucleic acid sequences from pancreatic tumors, and related
PT   proteins.
XX
PS   Claim 2; Page 272; 502pp; German.
XX
CC   This invention describes novel polypeptides and their encoding nucleic
CC   acids derived from human pancreatic tumor tissue which have cytostatic
CC   activity. The sequences are also useful in producing pharmaceutical
CC   compositions for treatment of pancreatic tumors. AA252858-Z53014
CC   represent expressed sequence tag (EST) fragments derived from a human
CC   pancreatic tumor cDNA library and which encode the proteins represented
CC   in AAY73814-Y74252
XX
SQ   Sequence 1376 BP; 294 A; 390 C; 448 G; 244 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   1376
Score: 1.18e-68   Length: 1376
Percent Similarity: 69.00%   Matches: 146
Best Local Similarity: 68.93%   Conservative: 47
Query Match: 52.14%   Mismatches: 83
DB: 47.73%   Indels: 4
                     2   Gaps: 3

US-09-889-733B-2 (1-285) x AAZ52967 (1-1376)

Qy   3 GlyLeuGlyHisPro---SerAlaPheGlyArgAlaThrHisAlaValAlaValArgAlaLeu 21
             ||||||:| ||||||||||||||||||| :::::: ||
Db   294 GGGATGGGGACCCCGGGGAGGGGCTGGGCCGGCTGCTCCCATGCCCGTCCCCGGGGAGTC 353

Qy   22 ProGluSerLeuCysGlnHisAlaAlaArgSerAlaLysGlyGluGluValAlaAspValAla 41
             ||||||||||||||||||| ||||||| ::::|| |::::: ||
Db   354 CCAGAGAGCCTGGCTCGGCGGGGAAGGTCGGGGGGCTGGCCTTCCCGCTCTGGATCTGGCC 413

Qy   42 ArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlyValLeuSerLeuGlyLeuGln 61
             ::::|||::||||| ||| || ::|||||||||||
Db   414 AAAGCTCAAAGGGAGCACGGGGTGCTGGGAGGTAAACTGAGGCAACGACTGGGGCTACAG 473

Qy   62 ValValGlyLeuProAlaAlaAspCysSerLeuProAspCysValPheValGluAspValAla 81
             :::::|||||||||| ::: ||| :|| ||||||||||::: ||||||||||
Db   474 CTGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCCGCTGCTTGGCGACACGGCC 533

Qy   82 ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlyGlu 101
             ::::::: ||| |||||||||||||| ::: |||||||||||| ||||||
Db   534 GTGATCCAAGGGGACACGGCCTAATCACGCGGGCCCTGGAGCCCCGCTCGTAGGCCAGAG 593

Qy   102 ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp 121
             :::::|||||||||||||||||| ::: |||||||||||| ::: ||||||
Db   594 GTCGATGGGATCCGGCAAAGCCCTGGCAAGACCCTGGGGCTCGAATTGTGGAAAATAGGAGAC 653

Qy   122 GluAsnAlaThrLeuAspGlyGlyAspGlyIleAlaGlyGluPheValLeuPheValGluPheValGly 141
             |||||||||||||||||| ::: |||||||||||||||||||||||||||
Db   654 GAGAACGCGCACGCTGGATGGCACTGACGTTCTCTTCACCGGCCGGGAGTTTTTCGTAGGC 713

Qy   142 LeuSerLysArgThrAsnGlnArgGlyIleAlaGluIleLeuAlaAspThrPheLysAspTyr 161
             ||||||||||||||||| :::| :::::: || ::: ||
Db   714 CTCTCCCAAATGACCAATCACCGAGGAGCTGAGATCGTGGCGACACGTTCCGGGACTTC 773

Qy   162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
             ||||||||||||| ::: |||||||||||||||||||||| :::
Db   774 GCCGTCCACTGTCCCAGTCTCGGGGTCCTCCACCTGCGCTGGGCATGGGG 833

Qy   182 GlyProAsnLeuLeuAlaIleGlySerSerGluSerAlaGlnIleAlaLeuLysIleMet 201
             ::: |||||||||||||||||||||| :::||
Db   834 GGACCTCGCACTGTTGTGCGAGGCAGCACGGCTGCCCAAAAGGCGTCCCGGGCAATG 893

AC   AAA99390;
XX
DT   22-JAN-2001 (first entry)
XX
DE   Human melanoma growth related factor-2 cDNA sequence.
XX
KW   Melanoma growth related factor-2; MGRF-2; human; ss.
XX
OS   Homo sapiens.
XX
PN   CN1257921-A.
XX
PD   28-JUN-2000.
XX
PF   18-DEC-1998;   98CN-0125527.
XX
PR   18-DEC-1998;   98CN-0125527.
XX
PA   (UYFU-) UNIV FUDAN.
XX
PI   Yu L, Fu Q, Zhang H;
XX
DR   WPI; 2000-544294/50.
DR   P-PSDB; AAB26802.
XX
PT   Human melanoma growth correlation factor and its coding sequence,
PT   preparing process and usage.
XX
PS   Claim 1; Page 14; 20pp; Chinese.
XX
CC   The present invention relates to a melanoma growth related factor (MGRF-
CC   2) cDNA and protein sequence, isolated from humans. The present sequence
CC   represents the cDNA of the invention
XX
SQ   Sequence 845 BP; 152 A; 250 C; 285 G; 158 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   3.56e-66   Length: 845
Score: 669.50   Matches: 141
Percent Similarity: 69.52%   Conservative: 46
Best Local Similarity: 52.42%   Mismatches: 78
Query Match: 46.05%   Indels: 4
DB: 3   Gaps: 2

US-09-889-733B-2 (1-285) x AAA99390 (1-845)

Qy   3   GlyLeuGlyHisPro---SerAlaPheGlyArgAlaThrHisAlaValValArgAlaLeu   21
        |||:||||||| |||||||:|::|:||||
Db   41   GGGATGGGACACGCCGGGGGAGGGGCTGGCCCGCTGCCTGCTCCCATGCCCTGATCCGGGGGAGTC   100

Qy   22   ProGluSerLeuCysGlnHisAlaAlaLeuArgSerAlaLysGlyGluValAspValAla   41
        ||||| ||:|||||:|||||
Db   101   CCAGAGAGCCTGGCGTCGGGGGAAGGTCGGGGCCTGGCCTTCCCGGCTCTGGATCCGCC   160

Qy   42   ArgAlaGluArgGlnHisGlnHisGlyLeuTyrValGlyValLeuGlyValLeuGlyLeuGln   61
        ||:||| |||||||| |||:|:
Db   161   AAAGCTCAAAGGGAGCACGGGGGTGCTGGGAGGTAAACTGAGGCAACGACTGGGGCTACAG   220

Qy   62   ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla   81
        :::|||||| :||:||||||
Db   221   CTGCTAGAACTGCCACCTGAGGAGTCAATTGCCGCTGGCACCGCTGCTTGGCCGACCACGCGGCC   280

Qy   82   ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu   101
        |:||| :::|||||
Db   281   GTGTGACCAAGGGGACACGGCCCTAATCACGCGGCCCGGCCCCGGCTAGGCCAGAG   340

Qy   102   ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp   121
        :::::|||||| ::::|:
Db   341   GTCGATGAGTCCGCCAAAGCCCTGCAAGACCTGGGGCTCCGAATTGTGGAAATAGGAGAC   400

Qy   122   GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGly   141
        ||||||||||| |||||:||
Db   401   GAGAACGGCACGCTGGATGGCACTGACGTTCTCTTCACCGGCCCGGGAGTTTTCGTAGGC   460

XX
SQ   Sequence 858 BP; 150 A; 268 C; 281 G; 159 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.87e-68   Length: 858
Score: 688.00   Matches: 145
Percent Similarity: 51.97%   Conservative: 47
Best Local Similarity: 47.32%   Mismatches: 83
DB: 3   Indels: 4
                                      Gaps: 3

US-09-889-733B-2 (1-285) x AAA47656 (1-858)

Qy   4   LeuGlyHisPro---SerAlaPheGlyArgAlaThrHisAlaValValArgAlaLeuPro   22
        :::|||||| :::|
Db   1   ATGGGACGCCGGGGGAGGGGCTGGCCCGCTGCCTGCTCCCATGCCCTGATCCGGGGGAGTCCCA   60

Qy   23   GluSerLeuCysGlnHisAlaAlaLeuArgSerAlaLysGlyGluValAspValAlaArg   42
        ||| :::|:|||
Db   61   GAGAGCCTGGCGTCGGGGGAAGGTGCGGCCTGGCCTTCCCGGCTCTGGATCTGGCCAAA   120

Qy   43   AlaGluArgGlnHisGlnHisGlyLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnVal   62
        ::|:||:| :::   :|:
Db   121   GCTCAAAGGGAGCACGGGGTGCTGGGAGGTAAACTGAGGCAACGACTGGGGCTACAGCTG   180

Qy   63   ValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaVal   82
        :::||| |:|
Db   181   CTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCGCTGCTTGGCGACACGGCCGTG   240

Qy   83   ValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluVal   102
        :::| :::||||||||||
Db   241   ATCCAAGGGGACACGGCCCTAATCACGCGGCCCGGCCCCTGGAGCCCCGCTCGTAGGCCAGAGTC   300

Qy   103   AspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGlu   122
        :::::|:| :::||
Db   301   GATGGAGTCCGCAAAGCCCTGCAAGACCTGGGGCTCCGAATTGTGGAAATAGGAGGACGAG   360

Qy   123   AsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeu   142
        ||||||||||| |:|
Db   361   AACCGCAGCCTGGATGGCACTGACGTTCTTCTTCACCGGCCGGGAGTTTTTCGTAGGCCTC   420

Qy   143   SerLysArgThrAsnGlnArgGlyAlaGlyIleLeuLeuAlaAspThrPheLysAspTyrAla   162
        ::::|||| :::::|:
Db   421   TCCAAATGGACCAATCACCGAGGAGTGAGATCGTGCCGGACACGTTCCGGGACTTCGCC   480

Qy   163   ValSerThrValProValAlaAspGlyLeuHisLeuCysSerPheCysSerMetAlaGly   182
        ::| ||| |:
Db   481   GTCTCCACTGTGCCAGTCTCGGGTCCTCCCACCTGCGCGGTCCTCTGCGCATGGGGGGA   540

Qy   183   ProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGln   202
        ::| :|||||||||||
Db   541   CCTCGCACTGTTGTGGCAGCAGCGACGGTCGCGGAAGGTCGGGGCCAATGGCA   600

Qy   203   GlnMetSerAspHisArgTyrAspLysProAspAspProAlaAlaAlaAsnCys   222
        ::::| |:::|||||:|
Db   601   GTGCTGACAGATCACCCATATGCCTCCCTGACCCCCAGATGACGCAGCTGCTGACTGT   660

Qy   223   IleTyrLeu-----AsnIleProAsnLysGlyHisValLeuLeuGluHisArgThrProGlu   240
        :::|| |||
Db   661   CTCTTTCTTCGTCCTCGGGTTGCCTGGTGTGCCCCCTTCCTCCTGCCACCGTGGAGGTGGG   720

Qy   241   GluTyrProGluSerAlaLysValTyrGluLysLeuGluLysAspHisMetLeuIleProVal   260
        ::| ||| :||
Db   721   GATCTGCCAAGGGGACACAGCCGCAGGAGGACGTTGCTGCAGAAGCTCTGATGTGCACCCTGGTACCTGTG   780

Qy   261   SerMetSerGluLeuGlyLysValAsp---GlyLeuLeuThrCysCysSerValLeu   278
        :::|| :::| ||
Db   781   TCCTGCTCAGAACTGAGAGAAGCTGCGCGCCGGGCTCCAGCTCCCCTGCCTCCTGGTGCTC   837

RESULT 15
AAA99390
ID   AAA99390 standard; cDNA; 845 BP.
XX

```
Qy   142 LeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyr 161
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||:::
Db   461 CTCTCCAAATGGACCAATCACCGAGAGACTGAGATCGTGGCGGACACGTTCCGGGACTTC 520

Qy   162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
          ||||||||||||||||||||||||||||:::   |||||||||    |||      |||
Db   521 GCCGTCTCCACTGTGCCAGTCTCGGGTCCTCCCACCTGCGGGGTCTCTGCCGGCATGGGG 580

Qy   182 GlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMet 201
          |||  ::  ||||||||||||||||||||||||| ||||:::::||||||
Db   581 GGACCTCGCACTGTTGTGGCAGGCAGCAGCGACCTGCCCAAAAGGCTGTCCGGGCAATG 640

Qy   202 GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsn 221
          :::::||||||||||  ||||||:::||||||||||||||||||||||||||:::
Db   641 GCAGTGCTGACAGATCACCCATATGCCTCCCTGACCCTCCCAGATGACGCAGCTGCTGAC 700

Qy   222 CysIleTyrLeu-----AsnIleProAsnLysGlyHisValLeuLeuHisValArgThrPro 239
          |:::::||||||:||         ||||||||||||||||||||
Db   701 TGTCTCTTTCTTCGTCCTGGGTGCCTGGTGTGCCCCTTTCCTCCGTGCACCGTGGAGGT 760

Qy   240 GluGluTyrProGluSerAlaLysVal-TyrGluLysLeuLysAspHisMetLeuIlePr 259
          |||     :::   ||  ||||||||   |||||||||  |||:::|||
Db   761 GGGGATCTGCCCAACAGCCAGGAGGCACCTGCAGAAGCTCTCTGATGTCACCCTGGTACC 820

Qy   259 oValSerMetSerGluGluLys 267
          ||||||  |||||||||||||
Db   821 TGTGTCCTGCTCAGAACTGGAGAAA 845
```

Search completed: June   7, 2004, 23:27:29
Job time : 387 secs

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: June 7, 2004, 23:12:01 ; Search time 2529 Seconds
(without alignments)
3365.252 Million cell updates/sec

Title: US-09-889-7333-2
Perfect score: 1454
Sequence: 1 MAGLGHPSAFGRATHAVVRA.........EKVDGLLTCCSVLINKKVDS 285

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27511289 seqs, 1491109276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09889733/runat_07062004_083528_7982/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09889733 @CGN_1_5180 @runat_07062004_083528_7982 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1389	95.5	1109	11	AK010430	AK010430 Mus muscu
	2	1217	83.7	1001	13	BY709864	BY709864 BY709864
	3	1127	77.5	746	14	CB249709	CB249709 UI-M-EX0-
	4	1125	77.4	862	13	BQ899468	BQ899468 AGENCOURT
	5	1118	76.9	889	12	BI763990	BI763990 603049879
	6	1077	74.1	760	14	CA512485	CA512485 UI-R-FJ0-
c	7	1064	73.2	822	9	AI647789	AI647789 uk43D05.x
	8	1056	72.6	656	13	BY742062	BY742062 BY742062
	9	975.5	67.1	860	14	CF548604	CF548604 AGENCOURT
c	10	956.5	65.8	800	9	AA986117	AA986117 uc81h06.x
	11	945	65.0	698	13	BY754930	BY754930 BY754930
	12	925	63.6	776	14	CD751679	CD751679 AGENCOURT
	13	896	61.6	731	10	BB611533	BB611533 BB611533
	14	892.5	61.4	897	14	CK019367	CK019367 AGENCOURT
	15	892	61.3	923	12	BI332932	BI332932 602982027
	16	881	60.6	522	9	AI751463	AI751463 cn10b04.x
	17	870	59.8	792	14	CD750695	CD750695 AGENCOURT
	18	867.5	59.7	777	14	CD349526	CD349526 UI-M-FY0-
	19	864	59.4	645	29	CG477727	CG477727 OST7734 M
	20	857	58.9	685	14	CB247154	CB247154 UI-M-FI0-
	21	853.5	58.7	802	14	CA88435	CA88435 hab63f06.
	22	822	56.5	501	29	CG617513	CG617513 OST311187
	23	820	56.4	702	13	BY731938	BY731938 BY731938
	24	816.5	56.2	1035	10	BF579996	BF579996 60205987
	25	806.5	55.5	788	14	CB248156	CB248156 UI-M-FI0-
	26	796	54.7	825	14	CD752232	CD752232 AGENCOURT
	27	794.5	54.6	654	9	AU296050	AU296050 AU296050
	28	789	54.3	722	14	CK142177	CK142177 AGENCOURT
	29	785	54.0	486	9	AI463821	AI463821 va31a06.y
	30	772	53.1	455	29	AY406329	AY406329 Homo sapi
	31	771	53.0	455	29	AY406330	AY406330 Pan trogl
	32	770.5	53.0	806	9	AU051351	AU051351 AU051351
	33	764	52.5	817	12	BI221174	BI221174 60293511
	34	756	52.0	681	12	BM963225	BM963225 UI-M-EQ0-
c	35	746	51.3	569	9	AI265716	AI265716 uj05e03.x
	36	744	51.2	556	12	BM964027	BM964027 UI-M-EQ0-
	37	741	51.0	455	29	AY406331	AY406331 Mus muscu
	38	718	49.4	432	10	BE864630	BE864630 UI-M-BH1-
	39	711.5	48.9	867	14	CD363134	CD363134 AGENCOURT
	40	708.5	48.7	783	14	CK138561	CK138561 AGENCOURT
	41	706.5	48.6	639	13	BQ399610	BQ399610 NISC_mp04
c	42	705	48.5	547	9	AI661738	AI661738 va31a06.x
c	43	701	48.2	624	10	AW514143	AW514143 hd24a02.x
c	44	697.5	48.0	776	14	CB317151	CB317151 AGENCOURT
	45	694	47.7	1351	11	AF070667	AF070667 Homo sapi

ALIGNMENTS

RESULT 1
AK010430
LOCUS   AK010430   1109 bp   mRNA   linear   HTC 20-SEP-2003
DEFINITION   Mus musculus ES cells cDNA, RIKEN full-length enriched library,
             clone:2410006N07 product:dimethylarginine dimethylaminohydrolase 1,
             full insert sequence.
ACCESSION   AK010430
VERSION   AK010430.1 GI:12845867
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE   1
AUTHORS   Carninci,P. and Hayashizaki,Y.
TITLE   High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED   10349636
REFERENCE   2
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED   11042159
REFERENCE   3
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED   11076861
REFERENCE   4
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 1109)
AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,Y., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT   Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAGGATCCAAGAGCTCTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTAATTAAATTAAATTCCCCCCCCCCCC 3']. cDNA was cleaved
            with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
            Host: SOLR.

FEATURES   Location/Qualifiers
     source   1..1109
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2410006N07"
                     /db_xref="MGI:1901307"
                     /db_xref="taxon:10090"
                     /clone="2410006N07"
                     /cell_type="ES cells"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
     CDS   137..994
                     /note="unnamed protein product; dimethylarginine
                     dimethylaminohydrolase 1 (MGD|MGI:1916469, GB|AK010430,
                     evidence: BLASTN, 100%, match=1109)
                     putative"
                     /codon_start=1
                     /protein_id="BAB26932.1"
                     /db_xref="GI:12845868"
                     /translation="MAGLGHPSAFGRATHAVVRAPPESLCRHALRRSQGEEVDPARAE
                     RQEELYVGLGSKLGLGQVVQLPADESLPDCVFVEDVAVVCEETALITRPGAPSRRKEV
                     DMMKEALEKLQLNIVEMKDENATLDGGDVLTGREFFVGLSKRTNQRGAEHLADTFKD
                     YAVSTVPVADSLHLKSPCSMAGPNLIAIGSSESAQKALKINQMSDHRYDKLTVPDDM
                     AANCIYLNIPSKGHVLLHRTPEYPBSAKVYEKLKDHLLIPVSNEMKVDGLLTCCS
                     VFINKKIDS"

ORIGIN
Alignment Scores:
Pred. No.:   1.11e-152   Length: 1109
Score: 1389.00   Matches: 269
Percent Similarity: 97.89%   Conservative: 10
Best Local Similarity: 94.39%   Mismatches: 6
Query Match: 95.53%   Indels: 0
DB: 11   Gaps: 0

US-09-889-733B-2 (1-285) x AK010430 (1-1109)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 ATGGCGGGCCTCGGCCACCCCTCGGCCTTCGGCCGGGCCACCCACGCCGTGGTGCGGGCT 196

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAspVal 40
            |||||||||||||||||||||||| ::::||||||||||||||||||||||
Db   197 CCGCCCGAGTCCCTGTGCCGCCACGGCCTGAGGCGGTCGCAGGGCGGAGAGGTGGATTTC 256

Qy   41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysGlyLeu 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   257 GCTCGCGCCGAGCGCCAGCACGAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTG 316

Qy   61 GlnValValGlnLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
            ||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db   317 CAGGTGGTGCAGCTGCCCGCCGACGAGAGCCTGCCCGACTGCGTGTTCGTGGAGGACGTG 376

Qy   81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   377 GCCGTCGTGTGCGAGGAGACGGCCCTCATCACCCGGCCCGGGGCCCCCAGCCGCCGCAAG 436

Qy   101 GluValAspMetMetLysGluAlaGluLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   437 GAGGTTGACATGATGAAAGAAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAA 496

Qy   121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuThrGlyArgGluPhePheVal 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   497 GATGAAAATGCCAACTTGGATGGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTTGTG 556

Qy   141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluHisLeuAlaAspThrPheLysAsp 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   557 GGCCTTTCCAAAAGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 616

Qy   161 TyrAlaValSerThrValProValAlaAspSerLeuHisLeuLysSerPheCysSerMet 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   617 TACGGCAGTCTCTACAGTCCCTGTGCCAGCATGCCTGGCATTTAAAGAGTTTCTGCAGCATG 676

Qy   181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
Db   677 GCCGGACCCAACCTGATTGCAATAGGGTCCAGCGAATCTGCACAGCAGAGGCCCTCAAGATC 736

Qy   201 MetGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db   737 ATGCAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGACATGGCCGCC 796

Qy   221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db   797 AACTGCATATATCTAAATATCCCAGCAAAGGGCATGTCTTGCTGCACCGGACCCCAGAA 856

Qy   241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
Db   857 GAGTACCCAGAAAGCGCAAAGGTCTATGAGAAACTCAAGGACCATACTACTGATCCCTGTG 915

Qy   261 SerMetSerGluLeuGluLeuValAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
Db   917 AGCAACTCGGAGATGGAAAAGGTGGACGGCTTGCTCACCTGTCTCCGTTTTTATTAAC 976

Qy   281 LysLysValAspSer 285
Db   977 AAGAAGATAGACTCC 991

RESULT 2
BY709864
LOCUS   BY709864   1001 bp   mRNA   linear   EST 16-DEC-2002
DEFINITION   BY709864 RIKEN full-length enriched, ES cells Mus musculus cDNA
             clone 2410006N07 5', mRNA sequence.
ACCESSION   BY709864
VERSION   BY709864.1 GI:27121074
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1001)
AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
             Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
             Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
             Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
             Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
             Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
             Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
             Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
             Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
             Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
             Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
             Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
             Maltais,L., Marchionni,L., Mckenzie,L., Miki,H., Nagashima,T.,
             Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
             Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
             Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
             Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
             Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
             Verardo,R., Wagner,L., Wahlested,C., Wang,Y., Watanabe,Y.,
             Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
             Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
             Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
             Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
             Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
             Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
             Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
             Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
             Laboratory for Genome Exploration Research Group, RIKEN Genomic
             Sciences Center (GSC), Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)

             1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
             Tel: 81-45-503-9222
             Fax: 81-45-503-9216
             Email: genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/
             Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
             Fukuda,S., Hashizume,W., Hayashida,K., Hirotani,T., Hori,F.,
             Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
             Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
             Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
             Sano,H., Sasaki,K., Sato,K., Shibata,K., Shiraki,T., Tagami,H.,
             Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
             Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
             Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new
             genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
             sequencing pipeline with 384 multicapillary sequencer. Genome Res.
             10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
             encyclopedia: real-time sequence clustering for construction of a
             nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
             further details.

FEATURES   Location/Qualifiers
     source   1..1001
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="2410006N07"
                     /cell_type="ES cells"
                     /lab_host="SOLR"
                     /clone_lib="RIKEN full-length enriched, ES cells"
                     /note="Site 1: XhoI; Site 2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAAGAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 5.0 and subtraction to Rot = 25.0. Second strand
                     cDNA was prepared with the primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAATCCCCCCCCCCCCCC 3']."

ORIGIN

US-09-889-733B-2 (1-285) x BY709864 (1-1001)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaaThrHisAlaValArgAla 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   134 ATGGCCGGCCTCGGCCACCCCTCGGCCTTCGGGCCGGGCCACCACGCCGCGTGGTGCGGGCT 193

Qy   21 LeuProGluSerLeuCysGlnHisHisAlaLeuArgSerAlaLysGlyGluGluGluValAspVal 40
           |||||||||||||::::::||||||||||||||||||||||||||||
Db   ...

```
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.

FEATURES             Location/Qualifiers
source               1..746
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5719830"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 15.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_EX0"
                     /note="Organ: brain; Vector: pYX-Asc; Site_1: ECOR I;
                     Site_2: Not I; The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured mRNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was size selected according to mRNA size fraction,
                     ligated with ECOR I adaptor, digested with Not I, and then
                     cloned directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA
                     tail, is GTGCGTGGAA. This library was created for the
                     University of Iowa Brain Molecular Anatomy Project
                     (BMAP), 'Gene Discovery in the Developing Mouse Nervous
                     System', supported by National Institutes of Mental Health
                     (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
Pred. No.:              4.88e-122      Length:          746
Score:                  1127.00        Matches:         219
Percent Similarity:     97.83%         Conservative:    6
Best Local Similarity:  95.22%         Mismatches:      5
Query Match:            77.51%         Indels:          0
DB:                     14             Gaps:            0

US-09-889-733B-2 (1-285) x CB249709 (1-746)

Qy   56  SerLysLeuGlyLeuGlnValValGluGluProAlaAspGluSerLeuProAspCysVal  75
         |||||||||||||||||||||||||||::|||||||||||||||||||||||||:||
Db    3  AGCAAGTCTGGGGCTGCAGGTGGTGCAGCTGCCGCCGGACGAGCCTGCCGACTGCGTG  62

Qy   76  PheValGluAspValAlaAlaValCysGluThrAlaLeuIleThrArgProGlyAla  95
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63  TTCGTGGAGGACGTGGCCGTCGTGTGCGAGGAGACGGCCCTCACCCCGCCCGGGGCG  122

Qy   96  ProSerArgArgLysGluValAspMetMetLysGluAlaLeuGluLysGlnLeuAsn  115
         |||||||||||||||||||||::|||||||||||||||||||||||||||||||||
Db  123  CCCAGCCGCCAGGAAGGAGGTTGACATGATGAAAGAGGCTTTGGAAAAACTTCAGTCAAC  182

Qy  116  IleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGly  135
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183  ATAGTAGAGATGAAAGATGAAAATGCAACTTTGGATGGTGGGGACGTCCTATTCACAGGC  242

Qy  136  ArgGluPhePheValGlyLeuSerLysArgThrAsnGluAlaArgGlyAlaGlyLeuLeuLeuAla  155
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  243  AGAGAATTTTTTGTGGGCCTTTCCAAAAGAACAAATCAACGAGGTGCTGAAATCTTGGCT  302

Qy  156  AspThrPheLysAspTyrArgAlaValSerThrValProValAlaAspGlyLeuHisLeuLys  175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  303  GATACTTTTAAGGACTACGCAGTCTTCTACAGTCCCTGGCCGATTCTTTGCATTTAAAG  362

Qy  176  SerPheCysSerMetAlaGlyTyrProAsnLeuIleAlaIleGlySerGluSerAlaGln  195
         ||||||||||||||||||||||||||||||||||||::||||||||||||||||||||
Db  363  AGTTTCTGCAGCATGGCCGGACCCAACCTGATTGCCAATAGGGTCCAGCGAATCTCCACAG  422

Qy  196  LysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValPro  215
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  194  CCGCCCAGTCCCTGTGCCGCACGGCCACGCTGAGCGGCGTCGAGGGAGGAGGTGGATTC  253
Qy   41  AlaArgArgAlaGluArgGlnHis-GlnLeuTyr-ValGlyValLeuGlySerLysLeuGlyL  60
         ||||||||||||||||||| ::||||||||| :::||||||||||||||||||||
Db  254  GCTCGCCGCGAGCGCCAGCACCAGCTCTACCGTGGGCGTGCTGGGCAGCAAGCTGGGGC  313
Qy   60  euGlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspV  80
         |||||||||||||||||||||||||||::|||||||||||||||||||||||||
Db  314  TGCAGGTGGTGCAGCTGCCGCCGGACGAGCCTGCCGACTGCGTGTTCGTGGAGGACG  373
Qy   80  alAlaValCysGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgArgL  100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  374  TGGCCGTCGTGTGCGAGGAGACGGCCCTCATCACCCGCCCCGGGGCGCCCAGCCGCCAGGA  433
Qy  100  ysGluValAspMetMetLysGluAlaLeuGluLysGlnLeuAsnIleValGluMetL  120
         ||||||||||||::||||||||||||||||||||||||||||||||||||
Db  434  AGGAGGTTGACATGATGAAAGAGGCTTTGGAAAAACTTCAGTCAACATAGTAGAGATGA  493
Qy  120  ysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheV  140
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  494  AAGATGAAAATGCAACTTTGGATGGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTG  553
Qy  140  alGlyLeuSerLysArgThrAsnGluAlaArgGlyAlaGlyLeuLeuLeuAlaAspThrPheLysA  160
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  554  TGGGCCTTTCCAAAAGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGG  613
Qy  160  spTyrArgAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerM  180
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  614  ACTACGCAGTCTTCTACAGTCCCTGGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCA  673
Qy  180  etAlaGlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysI  200
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  674  TGGCCGGACCCAACCTGATTGCCAATAGGGTCCAGCGAATCTCCACAGAAAGCGCTCAAGA  733
Qy  200  leMetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspIleAlaA  220
         |||||||||||||||||::|||||||||||||:||||||||||
Db  734  TCATGCAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGACATGGCCG  793
Qy  220  laAsnCysIleTyrLeuAsnIleProAsnIleGlyHisValLeuLeuHisArgThrProG  240
         ||||||||||||||::|||:..:::
Db  794  CCAACTGCATATATCTAAATATCCCCAGCATAGGGCATGCTTGCTGCACTGAACCCCAG  853
Qy  240  luGluTyrProGluSerAlaLysValTyrGluGluAspHisMetLeuIleProV  260
         ||||||||||||||||::|||||||
Db  854  AAGAGTACCCAGAAAGCGCAAACGTCTATGAGGAAGACCATCTACTGACTCCTCTG  913
Qy  260  alSerMetSerGluGluLysValAspGlyLeuLeuThrCysCysSerValLeuIle  279
         |||||||| :::|||: .::
Db  914  TGAGCAC-TCGGAGATGGAAAAGTGGACGGCC---GTCACCTGCGC-TCCGTTTTATTA  967
```

```
RESULT 3
CB249709
LOCUS         CB249709                746 bp   mRNA   linear   EST 15-JUL-2003
DEFINITION    UI-M-EX0-byl-1-07-0-UI.r1 NIH_BMAP_EX0 Mus musculus cDNA clone
              IMAGE:5719830 5', mRNA sequence.
ACCESSION     CB249709
VERSION       CB249709.1   GI:28389411
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
  ORGANISM    Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1  (bases 1 to 746)
  AUTHORS     Marra,M., Lin,J., University of Iowa
  TITLE       NIH-MGC http://mgc.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL     Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Dr. James Lin, Univeristy of Iowa
              cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
              cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: MGC clone distribution information can be
```

Qy 21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGlyValAspVal 40
Db 191 CCGGCCGAGTCCCTGTCCGCCACGCATCGGCTGCGAGGCGCTGAGGAGGAGGTGGATTTC 250

Qy 41 AlaArgAlaGluArgGlnHisGlnLeuValTyrValGlyValLeuGlySerLysLeuLeu 60
Db 251 GCTCGGCGCGAGCGCCAGCAGCTCACGTGCGGTGCTGGGCAGCAAGTGGGGCTG 310

Qy 61 GlnValValGluLeuProAlaSerGluSerLeuProAspCysValPheValGluAspVal 80
Db 311 CAGGTGGTGCAGCTGCCCGCCAGCGAGAGCCTGCCCGACTGCGTGTTCGTGGAGACGTG 370

Qy 81 AlaValValCysGlyGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
Db 371 GCCGTCGTGTGCGGAGAGACGGCCCTCATCACCCGGCCCGGGGCGCCCAGCCGCAGGAAG 430

Qy 101 GluValAspMetMetLysGluLeuGluLysGlnLeuAsnIleValGluMetLys 120
Db 431 GAGGTTGACATGATGAAGGAGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAA 490

Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
Db 491 GATGAAAATGCAACTTTGGATGGTGGGGACGTCCATTCACAGGCAGCAGAATTTTTGTG 550

Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyLysAlaGluIleLeuAlaAspThrPheLysAsp 160
Db 551 GGCCTTTCCAAAAGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGAC 610

Qy 161 TyrAlaValSerThrValProValAlaAspValProAspCysSerPheCysSerSerMet 180
Db 611 TACGCAGTCTTCACAGTCCCTGTGGCCGATTCTTGCATTAAAGAGTTTCTGCAGCATG 670

Qy 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerArgAlaGlnLysAlaLeuLysIle 200
Db 671 GCCGGACCCAACCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAAGCCCTCAAGATC 730

Qy 201 MetGlnGlnMetSerAspHisArgTyrAspAspLeuThrValProAspAspIleAlaAla 220
Db 731 ATGCAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGACATGGCCGCC 790

Qy 221 AsnCysIleTyrLeuAsnIle-ProAsnLysGlyHis-ValLeu-LeuHisArgThrPro 239
Db 791 CACTGCATATATCTAAATATCCCCAGCAAAGGGCATGTCTTGGCTGCACCGGACCCCC 850

Qy 240 GluGlu 241
Db 851 CAAAAG 856

RESULT 5
BI763990
LOCUS   BI763990   889 bp   mRNA   linear   EST 25-SEP-2001
DEFINITION 6030498 79F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5189970 5',
            mRNA sequence.
ACCESSION   BI763990
VERSION   BI763990.1 GI:15755568
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 889)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:

Db 423 AAGGCCCTCAAGATCATGCAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCC 482
Qy 216 AspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeu 235
Db 483 GACGACATGGCCGCCAACTGCATATATCTAAATATCCCCAGCAAAGGGCATGTCTTGCTG 542

Qy 236 HisArgThrProGluGlyTyrProGluSerAlaLysValTyrGluLysLysLysAspHis 255
Db 543 CACCGACCCCGAAGGATACCCAGAAAGCGCANAGGTCTATGAGAAACTCAAGAGACCAT 602

Qy 256 MetLeuIleProValSerMetSerGluLeuGluLysGlyValAspGlyLeuLeuThrCysCys 275
Db 603 CTACTGATCCCGTGAGCAACTCGGAGATGGAAAAGGTGGACGGCCTTGCTCACCTGCTGC 662

Qy 276 SerValLeuIleAsnLysLysValAspSer 285
Db 663 TCCGTTTTTATTAACAAGAAGATAGACTCC 692

RESULT 4
BQ899468
LOCUS   BQ899468   862 bp   mRNA   linear   EST 16-AUG-2002
DEFINITION AGENCOURT_8750357 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6332815
            5', mRNA sequence.
ACCESSION   BQ899468
VERSION   BQ899468.1 GI:22291482
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 862)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
            Ph.D.
            cDNA Library Preparation: ResGen, Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13791 row: a column: 08
            High quality sequence stop: 704.
FEATURES   Location/Qualifiers
     source   1..862
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6332815"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC 130"
                     /note="Organ: oocysts; Vector: pCMV-SPORT6.1; Site_1:
                     EcoRV; Site_2: NotI; Cloned unidirectionally. Primer:
                     Oligo dT. Average insert size 1.95 kb. Constructed by
                     ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:   1.07e-121   Length: 862
Score: 1125.00   Matches: 226
Percent Similarity: 96.72%   Conservative: 10
Best Local Similarity: 92.62%   Mismatches: 5
Query Match: 77.37%   Indels: 3
DB: 13   Gaps: 0

US-09-889-733B-2 (1-285) x BQ899468 (1-862)

Qy 1 MetAlaAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValAlaArgAla 20
Db 131 ATGGCCGGCCTCGGCCACCCCTCCGGCCTTCGGCCGGCCACCCACGCCGTGGCGGGCT 190

http://image.llnl.gov
Plate: LLAM11474 row: n column: 19
High quality sequence start: 3
High quality sequence stop: 847.

FEATURES   Location/Qualifiers
  source   1..889
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5189970"
                  /lab_host="DH10B"
                  /clone_lib="NIH_MGC 116"
                  /note="Organ: pooled colon, kidney, stomach; Vector:
                  pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                  source anonymous pool of 3 colons, age 26 yo male, 49 yo
                  female, 71 yo male colon; 46 yo male kidney, and pool of 2
                  stomachs, 62 yo male and 70 yo female. Library is
                  oligo-dT primed and directionally cloned (EcoRV site is
                  destroyed upon cloning). Average insert size 1.4 kb,
                  insert size range 1-3 kb. Library is normalized and
                  enriched for full-length clones and was constructed by C.
                  Gruber (Invitrogen). Research Genetics tracking code
                  023. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.:   889
Score: 7.53e-121   Length: 889
                  1118.00   Matches: 237
Percent Similarity: 92.45%   Conservative: 8
Best Local Similarity: 89.43% Mismatches: 14
Query Match: 76.89%   Indels: 6
DB: 12   Gaps: 2

US-09-889-733B-2 (1-285) x BI763990 (1-889)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
            ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db   102 ATGGCCGGGCTCGGCCACCCCAGCGCCTTCGGCCGGGCCACCCACGCCGTGGTGCGGGCG 161

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAspVal 40
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   162 CTACCCGAGTCGCTCGGCCAGCACGCGCTGAGAAGGGCCAAGGGCGAGGAGGTGACGCGTC 221

Qy   41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlyValSerLysGlyLeu 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   222 GCCCGCGCCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCGTCGCAAGCTGGGGCTG 281

Qy   61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   282 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 341

Qy   81 AlaValValCysGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   342 GCCGTGGTGTGCGAGACGGCGCCCTCATCACCCGACCCCGGGGCGCCGAGCCGGAGGAAG 401

Qy   101 GluValAspMetMetLysGluValAlaLeuGluLysGlnLeuAsnIleValGluMetLys 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   402 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 461

Qy   121 AspGluAsnAlaThrLeuAspGlyArgTyrGlyValPheThrGlyArgGluPhePheVal 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   462 GATGAAAATGCAACTTTAGATGGCCGAGATGTTTATTCACAGGCAGAGAATTTTTTTGTG 521

Qy   141 GlyLeuSerLysArgThrAsnGlnArgGlnIleAlaGluIleLeuAlaAspThrPheLysLys 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   522 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTCGATACTTTTAAGGAC 581

Qy   161 TyrAlaValSerThrValProValAlaAspLysAlaHisLeuHisLysLysSer-PheCysSerSerMe 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   582 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTTCTGCAGCAT 641

Qy   180 tAlaGlyProAsnLeuIleAlaIleGlySerGluSerGluSerGlnAlaGlnAsnLysLysLeuLysIl 200
            |||||||||||||||||||||||||||||||||

located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CATCCTACT. This library
was created for the University of Iowa Program for Rat
Gene Discovery and Mapping (Val Sheffield, Bento Soares
and Tom Casavant)"

ORIGIN

Alignment Scores:
Pred. No.:   4.01e-116
Score: 1077.00   Length: 760
Percent Similarity: 98.16%   Matches: 209
Best Local Similarity: 96.31%   Conservative: 4
Query Match: 74.07%   Mismatches: 4
DB: Indels: 0
                                         Gaps: 0

US-09-889-733B-2 (1-285) x CA512485 (1-760)

Qy   69 GluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGluThrAla   88
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 GAGAGCCTGCCTGACTGCGTGTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAGACCGCC   61

Qy   89 LeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLysGluAla   108
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 CTCATCACCCGGCCCGGGGCGCCCTAGCCCGCAGGAGGAGGTTGACATGATGAAAGAGGCT   121

Qy   109 LeuGluLysLeuGlnIleuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGly   128
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 TTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATCCAACCTTAGATGGT   181

Qy   129 GlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThrAsnGln   148
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 GGGGACGTCCTATTCACAGGCAGAGAGTTTTTGTGGGCCTTTCCAAAAGACAAATCAA   241

Qy   149 ArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrValProVal   168
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 CGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGACTACGCAGTTTCCACAGTCCCCGTG   301

Qy   169 AlaAspGlyLeuHisIleuLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIle   188
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 GCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCATGGCTGGCCCCAACCTGATCGCAATA   361

Qy   189 GlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisArg   208
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 GGGTCCAGTGAATCTGCGCAGAAGGCCCTCAAGATCATGCAACAGAGTGACCACCGT   421

Qy   209 TyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIlePro   228
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 TATGACAAGCTCACTGTACCCGACGACGACATGGCGCCAACTGTATATATTTAAATATCCCC   481

Qy   229 AsnLysGlyHisValLeuuLeuHisArgThrProGluGluTyrProGluSerAlaLysVal   248
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 AGCAAAGGGCATGTCTTGCTGCACCGAACCCCAGAAGAGTACCCAGAAAGCGCAAAGGTT   541

Qy   249 TyrGluuLysLeuLysAspHisMetLeuIleLeuProValSerMetSerGluLeuGluLysVal   268
          |||||||||||||||::|::||||||||||||||||||||||||||||||||||||||||
Db   542 TATGAGAAGCTCAAGGACCATCTACTGATCCTGGTGAGCAATTCTGAGATGGAAAAGGTG   601

Qy   269 AspGlyLeuLeuThrCysCysSerValLeuIleuAsnLysLysValAspSer   285
          |||||||||||||||:||:::||||||||||||:::||||||||||||||||
Db   602 GACGGCTTGCTCACCTGCTGCTCCGTGTTTTTATTAACAAGAGACACAGACTCT   652

RESULT 7
AI647789/c
LOCUS   AI647789   822 bp   mRNA   linear   EST 30-APR-1999
DEFINITION uk43b05.x1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1971729 3' similar to TR:O08557 O08557
            N-G,N-G-DIMETHYLARGININE DIMETHYLAMINOHYDROLASE .; mRNA sequence.
ACCESSION   AI647789
VERSION   AI647789.1   GI:4726467
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS   1 (bases 1 to 822)
            Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
TITLE   The WashU-NCI Mouse EST Project 1999
JOURNAL   Unpublished (1999)
COMMENT   Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:988469
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end
            Possible reversed clone: similarity on wrong strand
            Seq primer: custom primer used
            High quality sequence stop: 516.
FEATURES   Location/Qualifiers
     source   1..822
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1971729"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse kidney mkia"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCGTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCGTCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."

ORIGIN

Alignment Scores:
Pred. No.:   1.56e-114
Score: 1064.00   Length: 822
Percent Similarity: 95.32%   Matches: 217
Best Local Similarity: 92.34%   Conservative: 7
Query Match: 73.18%   Mismatches: 10
DB: Indels: 3
                                         Gaps: 0

US-09-889-733B-2 (1-285) x AI647789 (1-822)

Qy   52 GlyValLeuGlySerLysLeuGlyLeuGlnValValLeuGluLeuProAlaAspGluSerLeu   71
          ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db   822 GGCGTGCTGGGCAGCAAGCTGGGCAGCTGCAGGTGGTGCAGCTGCCCCGGACGAGAGCCTG   763

Qy   72 ProAspCysValPheValGluAspValAlaValValCysGluGluThrAlaLeuIleThr   91
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   762 CCCGACTGCGTGTTCGTGGAGGACGTGGCCGTGGTGTGCGA-GAGACGGGCCTTCATCACC   704

Qy   92 Arg-ProGlyAlaProSerArgArgLysGluValAspMetMetLysGluAlaLeuGluLyy   111
          ||||:||||||||||||||||||::::::|||||||||||||||||||||||||||||
Db   703 CGCCCCGGGGGCGCCAGCCNGCAGCAAGAGGAGGTTGACATGATGAAAGAGGCTTTGGA-AA   645

Qy   111 sLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyValAspVa   131
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   644 ACTTCAGCTCAACATAGTAGAGATGANAGATGAAAAATGCAACTTGGATGGTGGGGACGT 585
          |||||||||||
Qy   131 lLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThrArgAsnGlnArgGlyAl 151

Db   584 CCTATTCACAGGCAGCAGAGAATTTNNTGCGGCCTTTCCAAAGACAAATCAACGAGGTGC 525

Qy   151 aGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrValProValAlaAspGl 171
Db   524 TGAAATCTTGGCTGATACTTTTAAGGACTACACGCAGTCTCTCACAGTCCCTGTGGCCGATTC 465

Qy   171 yLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSe 191
Db   464 TTTGCATTAAAAGAGTTTCTGCAGCATGGCCGGACCACCTGATTGCAATAGGGTCCAG 405

Qy   191 rGluSerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisArgTyrAspPly 211
Db   404 CGAATCTGCCAGAGGCCCTCAAGATCATGCAACAGATGAGTGACCATCGTTATGACCA 345

Qy   211 sLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGl 231
Db   344 GCTCACTGTACCCGACGACATGGCCGCCAACTGCATATATCCCCAGCAAAGG 285

Qy   231 yHisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluly 251
Db   284 GCATGTCTTGTGCACCGAACCCCAGAAGAGTACCCAGAAAAGCCCAAAGGTCTCATGAGAA 225

Qy   251 sLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValLysValAspGlyLe 271
Db   224 ACTCAAGGACCACTCTACTGATCCCTGTGAGCAACTCGGAGATGGAAAAGGTGGGACGGCTT 165

Qy   271 uIleuThrCysCysSerValLeuIleLeuAsnLysLysLysValAspSer 285
Db   164 GCTCACCTGCTGCTCCGTTTTTATTAACAAGAAGATAGACTCC 122

RESULT 8
BY742062
LOCUS   BY742062   656 bp   mRNA   linear   EST 17-DEC-2002
DEFINITION BY742062 RIKEN full-length enriched, 2 days neonate sympathetic
            ganglion Mus musculus cDNA clone 7120491N07 5', mRNA sequence.
ACCESSION   BY742062
VERSION   BY742062.1 GI:27167348
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 656)
AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,N., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasaw,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahleithen,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            RNA was provided by Dr. Akira Nakagawara (Div. of Biochemistry,
            Chiba Cancer Center Research Institute, 666-2 Nitona, Chuoh-ku,
            Chiba, 260-8717 Japan) whose assistance is gratefully acknowledged.
            Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES   Location/Qualifiers
     source   1..656
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="7120491N07"
                     /sex="mixed"
                     /tissue_type="sympathetic ganglion"
                     /dev_stage="2 days neonate"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 2 days neonate
                     sympathetic ganglion"
                     /note="Site 1: SalI; Site 2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5',
                     GAGAGAGAAGGAGATCCAAGAGCTCGTTTTTTTTTTTTTTTVN 3']. cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5',
                     GAGAGAGGATTCTCGAGTTAATTAATAATCCCCCCCCCCCC 3']. cDNA
                     was cleaved with XhoI and BamHI. Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC I.
                     Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES   Location/Qualifiers
     source   1..860
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:7018702"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NICHD_XGC_Brn1"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.5 kb. Constructed by Life
                     Technologies. Note: This is a Xenopus Gene Collection
                     (XGC) library."

ORIGIN

Alignment Scores:
Pred. No.:   Length: 860
Score: 4.7e-104   Matches: 201
Percent Similarity: 975.50   Conservative: 24
Best Local Similarity: 86.87%   Mismatches: 32
Query Match: 77.61%   Indels: 4
DB: 67.09%   Gaps: 1
                       14

US-09-889-733B-2 (1-285) x CF548604 (1-860)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
        ||||||||||||||||||||||||||||||||||| :|||||||||::|||:||
Db   77 ATGGCTGGACTAGCT---CTGTCAGGCTTTGGGAGATGCACCCACGTTATAGTTCGTGCA 133

Qy   21 LeuProGluSerLeuCysGlnHisAlaAlaLeuArgSerAlaLysGlyValGluValAspVal 40
        :::||||||:::||||||||| :|||||||||:: :||||:||::|:::||||:::
Db 134 ATTCCTGACTCTTTATGTCAAGAAGCTCTGACATAGAGAAGGCAGTGTGTGGATATC 193

Qy   41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
        |||| |||::::|| :::::|| ||||||||| ::|||||
Db 194 TCCCTGGCTAGGAAACAACATGAGCAGTATGGTGACGTGCTGAAGATAAACTGGGGCTG 253

Qy   61 GlnValValLeuProAlaAspCysValPheValGluAspVal 80
        |||||| :|:| |||||||::||||||||||||:::
Db 254 AGTGTGATAGAGCTTCAGCAGCGATGAATCCCTCCCAGTAGCGGCCCTTTGTGGAAGATGTG 313

Qy   81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyValAlaProSerArgArgLys 100
        |||| :||||||||||||||||||||||||||||||||
Db 314 GCTGTTCGTGTGTGAGGAAAACTGCCCTGATTACCAGACCTGGGCACCTAGCAGGAGGAGAAA 373

Qy 101 GluValAspMetMetLysGluValAlaLeuGluLysGlnLeuAsnIleValGluMetLys 120
        ||| ::: :||||| |||||||||| |||:|||||||
Db 374 GAGACTGAAGCTGTGAAGAAACACTGCAGGATCTTCAGCTAAATGGTGGTTGAAATGACA 433

Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
        |||||| |||:: :||::|||||||||||::|||||||||:::
Db 434 GATGAAAATGCTACATTAGATGGAGGAGATGTTTTGTTTACRGGAAAGGAATTCTTTGTG 493

Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGlyIleLeuLeuAlaAspThrPheLysAsp 160

~RNA was provided by Akira Nakagawara, Div. of
Biochemistry, Chiba Cancer Center Research Institute,
666-2 Nitona, Chuoh-ku, Chiba, 260-8717 Japan, whose
assistance we gratefully acknowledge."

ORIGIN

Alignment Scores:
Pred. No.:   Length: 656
Score: 9.45e-114   Matches: 209
Percent Similarity: 1056.00   Conservative: 6
Best Local Similarity: 98.62%   Mismatches: 3
Query Match: 95.87%   Indels: 1
DB: 72.63%   Gaps: 0
                       13

US-09-889-733B-2 (1-285) x BY742062 (1-656)

Qy   62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
        ||||||||||||||||::|||| :|||||||||||||||||||||||||
Db   3 GTGGTGCAGCTGCCCGCCGACGAGAGCCTGCCCGACTGCGTGTTCGTGGAGGACGTGGCC 62

Qy   82 ValValCysGluGluThrAlaLeuIleThrArgProGlyValAlaProSerArgArgLysGlu 101
        ||||||||||||||||||||||||||||||||||||||||
Db   63 GTCGTGTGCGAGGAGAGACGGCCCTTCATCACCCGGCCCGGGGCCGCCAGCCGCAGGAAGGAG 122

Qy 102 ValAspMetMetLysGluValAlaLeuGluLysGlnLeuAsnIleValGluMetLysAsp 121
        ||||||||||||||||||||||||||||||||||||||
Db 123 GTTGACATGATGAAAGAGGCTTTGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGAT 182

Qy 122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGly 141
        |||||||||||||||||||||||||||||||||||||||
Db 183 GAAAATGCAACTTTGGATGGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTGTGGGC 242

Qy 142 LeuSerLysArgThrAsnGlnArgGlyAlaGlyIleLeuLeuAlaAspThrPheLysAspTyr 161
        ||||||||||||||||||::||||||||||||||||:::
Db 243 CTTTCCAAAAGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTAC 302

Qy 162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
        ||||||||:| :||||||||||||||||||||| :::
Db 303 GCAGTCTCTCACAGTCCCGTGGCCGATTCTTCGCATTTAAAGAGTTTCTGCAGCATCGGCC 362

Qy 182 GlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysLeuAlaLeuLysLysIleMet 201
        ||||||||||||||||||||||||||||||||||||||||
Db 363 GGACCCAACCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAAGGCCCTCAAGATCATG 422

Qy 202 GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsn 221
        |||||||||||| ::::||||||||:::|||||||:::
Db 423 CAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGATGGCCGCCAAC 482

Qy 222 CysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGluGlu 241
        ||||||||::|||||||::||||||||||||||:::
Db 483 TGCATATATCTAAATATCCCAGCAAAGGGCATGTCTTGCTGCACCG-ACCCCAGAGAG 541

Qy 242 TyrProGluSerAlaLysValTyrGluLysMetHisMetLeuIleProValSer 261
        ||||||||| ||||:::||:::|||||||:::
Db 542 TACCCAGAAGCGCAAAGGTCTATGAGAAAACTCAAGGACCATCTACTGATCCCTGTGAGC 601

Qy 262 MetSerGluGluLeuGlyLysValAspGlyTyrLeuLeuThrCysSerValLeuIle 279
        ||||| :::||||||| |||:::||:::
Db 602 AACTGGGAATGGAAAAGTGGACGGCTTGCTCACCTGCTCCGTTTATTA 655

RESULT 9
CF548604
LOCUS   CF548604   860 bp   mRNA   linear   EST 22-SEP-2003
DEFINITION AGENCOURT 15597764 NICHD_XGC_Brn1 Xenopus laevis cDNA clone
            IMAGE:7018702 5', mRNA sequence.
ACCESSION   CF548604
VERSION   CF548604.1 GI:34885536
KEYWORDS   EST.
SOURCE   Xenopus laevis (African clawed frog)
  ORGANISM Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.

and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end
primer CGACCTGCAGCTCGAGCACA."

ORIGIN

Alignment Scores:
Pred. No.:   7.2e-102   Length: 800
Score: 956.50   Matches: 198
Percent Similarity: 92.00%   Conservative: 9
Best Local Similarity: 88.00%   Mismatches: 13
Query Match: 65.78%   Indels: 5
DB: 9   Gaps: 1

US-09-889-733B-2 (1-285) x AA986117 (1-800)

```
Qy    66  ProAlaAspGluSerLeuProAspCys-ValPheValGluAsp-ValAlaValValCysG  85
              |||||||||||||||||||||||||   ||||||||||||   |||||||||||||
Db   800  CCCGCCGACGAGAGCCTGCCCGACTGGCGTGTTTGTTGAAGAACGTGCCCGTCGTGTGCG  741

Qy    85  luGluThrArgLeu--IleThrArgProGlyAlaProSerArgArgLysGlu-ValAsp  103
              |||||||||        ||||||||||||||||||||||||:::||||| ||||||
Db   740  AGGAGACCCGGCTTCATCACCCGCCCGGGGCCCCAGCCCGCAAGAAGAGGGTTGCC  681

Qy   104  MetMetLysGluAlaLeuGlyLysGlnLeuAsnIleValGluMetLysAspGluAsn  123
              |||||||||||||||:|||||||||||:||||||||||||||||||||||||||
Db   680  ATGATGAAAGAGGCTTTGAAAAACTTCAGCTCACATAGTAGAGATGAAAGATGAAAAT  621

Qy   124  AlaThrLeuAspGlyGlyAspValLeuProGlu-PhePheValGlyLeuSe  143
              |||:|||||||||||||||||||||||||| |||||||||||||:||||
Db   620  GCANCTTTGGATGGTGGGGACGTCCTATCCACAGGCAGAGAATTTTNTGGTGGGCCTTTC  561

Qy   143  rLysArgThrAsnGlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysAspTyrAlaVa  163
              |||||||:|||||||||||||||||||:::||||||||:::||||||||||||||
Db   560  CAAAAGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTACGCAGT  501

Qy   163  lSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerSerMetAlaGlyPr  183
              |||||||:||||||||||||||||||||||||||||||:||||||||||||||||
Db   500  CTCTACAGTCCCTGTGGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCATGGCCGGACC  441

Qy   183  oAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGl  203
              |||||||||||||||||:||||||||||||||||||||:||||||||:|||||||||
Db   440  CAACCTGATCGCCAATAGGGTCCAGCGAATCTGCACCAGAAGGCCCTCCAAGATCATGCAACA  381

Qy   203  nMetSerAspHisArgThrValProAspAspIleAlaAlaAsnCysIl  223
              ||||||||:|||||||||||:|||||||||||||||||||||:|||
Db   380  GATGAGTGACCATCGTATGACAAGCTCACTGTACCCGACACATGGCCGCCAACTGCAT  321

Qy   223  eTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGluTyrPr  243
              |||||||||||||||||:|||||||||||||||||||||||:||||||:|||||
Db   320  ATATTAAATATCCCCAGCAAAGGGCATGTCTTGCTGCACCGAACCCCAGAGAGTACCC  261

Qy   243  oGluSerAlaLysValTyrGluLysLysAspHisMetLeuIleProValSerMetSe  263
              |||||||||||||||||||||||||:::||||||||:|||||||||||||||
Db   260  AGAAAGCGCAAAGTCTATGAGAAAACTCAAGAACCATCTACTGACCCCTGTGAGCAACTC  201

Qy   263  rGluLeuGlyLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysVa  283
              |||||||||:||||||||||||||:::||||||||||||:||||||||||||||:
Db   200  GGAGATGGAAAAGGTGCGACGGCTTGCTCACCTGCTCCGTTTTTATTAACAAGAAGAT  141

Qy   283  lAspSer  285
              |||||
Db   140  AGACTCC  134
```

RESULT 10
AA986117/c
LOCUS   AA986117   800 bp   mRNA   linear   EST 28-MAY-1998
DEFINITION uc81h06.x1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1432091 3' similar to TR:O08557 O08557
            N-G,N-G-DIMETHYLARGININE DIMETHYLAMINOHYDROLASE. ;, mRNA sequence.
ACCESSION   AA986117
VERSION   AA986117.1   GI:3167992
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 800)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:916159
            Possible reversed clone: similarity on wrong strand
            Seq primer: primer name ambiguous
            High quality sequence stop: 488.
FEATURES   Location/Qualifiers
     source   1..800
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1432091"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse kidney mkia"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested

```
Db   494  GGTTGTCGAAAGAACCACCAATCAACGTGGTGCTGAAATTCTGGCAGATACCTTTAAGGAT  553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   161  TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet  180
              ||||||||||||||||  :::  ||||||||||||||||||||||||||||||||
Db   554  TATGCAGTTTCCACTGTGCCAGTGTGCGAGACANTTGCACTTGAAGAGTTTCTGCAGCATG  613
Qy   181  AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle  200
              ||||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db   614  GCTGGACCAAACCTCATTGCCATTGGCTCGAGTGAAGCTGCCCAAAAAGCTCTAAAGACC  673
Qy   201  MetGlnGlnMetSerAspHisArgThrValProAspAspIleAlaAla  220
              ||||||||||||||||||||||||||:::|||||||||||||||:::||
Db   674  ATGCAGCAGATGAGTGATCACCCTACGACACAGCTGACATTGCCTGATAATGCTGCAGCT  733
Qy   221  AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu  240
              ||||||||:::||  ||||:::||||||||||||||||||||||||||:||||||
Db   734  AACTGCATTATT-AACATTCCAAGCAAAGGGCATGTTCTACTGCATCTTGCAACTGAG  792
Qy   241  GluTyrProGluSerAlaLysValTyrGlu-LysLeuLysLeuAspHisMetLeuIle  258
              |||||||||||||||||||||||||||| :::||  ||||||||||||:::
Db   793  AT-TTTTCAGAGAGTGCCAAAGGTGTTTGAAAAAACTGAAGAACACCATGCTGGTT  846
```

BY754930
RESULT 11
BY754930
LOCUS   BY754930   698 bp   mRNA   linear   EST 17-DEC-2002
DEFINITION BY754930 RIKEN full-length enriched, 8 cells embryo Mus musculus
            cDNA clone E860038F15 3', mRNA sequence.

ACCESSION　　BY754930
VERSION　　　BY754930.1　GI:27187153
KEYWORDS　　EST.
SOURCE　　　Mus musculus (house mouse)
ORGANISM　　Mus musculus
　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE　　1 (bases 1 to 698)
AUTHORS　　Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
　　　Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
　　　Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
　　　Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
　　　Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
　　　Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
　　　Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
　　　Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
　　　Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
　　　Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
　　　Kawaji,H., Kawasaki,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
　　　Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
　　　Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
　　　Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
　　　Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
　　　Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
　　　Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
　　　Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
　　　Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
　　　Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
　　　Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
　　　Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
　　　Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
　　　Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
　　　Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
　　　Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
　　　Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE　　　Analysis of the mouse transcriptome based on functional annotation
　　　of 60,770 full-length cDNAs
JOURNAL　　Nature 420, 563-573 (2002)
MEDLINE　　22354683
PUBMED　　12466851
COMMENT　　Contact: Yoshihide Hayashizaki
　　　Laboratory for Genome Exploration Research Group, RIKEN Genomic
　　　Sciences Center(GSC), Yokohama Institute
　　　The Institute of Physical and Chemical Research (RIKEN)
　　　1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
　　　Tel: 81-45-503-9222
　　　Fax: 81-45-503-9216
　　　Email: genome-res@gsc.riken.go.jp
　　　URL:http://genome.gsc.riken.go.jp/
　　　Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
　　　Fukuda,S., Hashizume,W., Hayashida,Y., Hirozane,T., Hori,F.,
　　　Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
　　　Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
　　　Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,N., Sakazume,N.,
　　　Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
　　　Takeda,H., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
　　　Direct Submission
　　　Computational Analysis of Full-Length Mouse cDNAs Compared with
　　　Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
　　　Normalization and subtraction of cap-trapper-selected cDNAs to
　　　prepare full-length cDNA libraries for rapid discovery of new
　　　genes. Genome Res. 10 (10), 1617-1630 (2000)
　　　RIKEN integrated sequence analysis (RISA) system--384-format
　　　sequencing pipeline with 384 multicapillary sequencer. Genome Res.
　　　10 (11), 1757-1771 (2000)
　　　Computer-based methods for the mouse full-length cDNA
　　　encyclopedia: real-time sequence clustering for construction of a
　　　nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
　　　cDNA library was prepared and sequenced in Mouse Genome
　　　Encyclopedia Project of Genome Exploration Research Group in Riken
　　　Genomic Sciences Center and Genome Science Laboratory in RIKEN.
　　　Division of Experimental Animal Research in Riken contributed to
　　　prepare mouse tissues.

Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
FEATURES　　Location/Qualifiers
　source　　1..698
　　　/organism="Mus musculus"
　　　/mol_type="mRNA"
　　　/strain="C57BL/6J"
　　　/db_xref="taxon:10090"
　　　/clone="B860038F15"
　　　/cell_type="8 cells"
　　　/dev_stage="8 cells embryo"
　　　/clone_lib="RIKEN full-length enriched, 8 cells embryo"

ORIGIN

Alignment Scores:
Pred. No.:　　　　　1..698　　　　　　Length:　　　　698
Score:　　　　　1.32e-100　　　　Matches:　　　　183
Percent Similarity:　945.00　　　　Conservative:　　5
Best Local Similarity:　97.41%　　　Mismatches:　　5
Query Match:　　　94.82%　　　Indels:　　　　0
DB:　　　　　　64.99%　　　　Gaps:　　　　　0
　　　　　　　13

US-09-889-733B-2 (1-285) x BY754930 (1-698)

```
Qy    93 ProGlyAlaProSerArgArgCysGluValAspMetMetLysGluAlaLeuGluLysLeu 112
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2 CCCGGGGCCCCAGCCCGCAGGAGAGAGTTGACATGATGAAAGAGGCTTTGAAAAACTT 61

Qy   113 GlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeu 132
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    62 CAGCTCAACATAGTAGAGATGAAAGATGAAAATGCAACTTTGGATGGTGNGGACGTCCTA 121

Qy   133 PheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGlu 152
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 TTCACAGGCAGAGAATTTTTGTGGGCTTTCCAAAAGAACAAATCAACGAGGTGCTGAA 181

Qy   153 IleLeuAlaAspThrPheLysAspTyrAlaValSerThrValProValAlaAspGlyLeu 172
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 ATCTTGGCTGATACTTTTAAGGACTACGCAGTCTCTACAGTCCCTGTGGCCGATTCTTTG 241

Qy   173 HisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGlu 192
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 CATTTAAAGAGTTTCTGCAGCATGGCCGACCCAACCTGATTGCAATAGGGTCCAGCGAA 301

Qy   193 SerAlaGlnLysAlaLeuLysIleMetGlnMetAspHisArgTyrAspLysLysLeu 212
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 TCTGCACAGAAGGCCCTCAAGATCATGCAACAGATGAGTGACCATCGTTATGACAAGCTC 361

Qy   213 ThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHis 232
         ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||:
Db   362 ACTGTACCCGACGACACATGGCCGCCAACTGCATATATCTAAATATCCCAGCCAAGGGCAT 421

Qy   233 ValLeuLeuHisArgThrProGluGlyTyrProGluSerAlaLysValTyrGluLysLeu 252
         |||||||::||||||||:|||||||||||||||||||||||||:|||||||||||
Db   422 GTCTTGCTGCACCGAACCCCAGAAGGCGCAGAAGGCTCATGAGAAACTC 481

Qy   253 LysAspHisMetLeuIleProValSerMetSerGluGluLysValAspGlyLeuLeu 272
         ||||||||:::|||:|||||:||:||||||||||||::::::||:::|||||:
Db   482 AAGGACCATCTACTGATCCCTGTGAGCAACTCGGAGATGGAAAAGGTGGACGGCTTGCTC 541

Qy   273 ThrCysCysSerValLeuIleAsnLysLysValAspSer 285
         |||||||||||||||||||||||||||||||||||
Db   542 ACCTGCTGCTCCGTTTTATTAACAAGAAGATAGACTCC 580
```

RESULT 12
CD751679
LOCUS　　　CD751679　　　　　　776 bp　　　mRNA　　linear　EST 30-JUN-2003
DEFINITION　AGENCOURT 14624576 NCI_CGAP_ZEmb2 Danio rerio cDNA clone
　　　IMAGE:6964621 5', mRNA sequence.
ACCESSION　CD751679
VERSION　　CD751679.1　GI:32335966
KEYWORDS　　EST.
SOURCE　　　Danio rerio (zebrafish)

```
ORGANISM  Danio rerio
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
          Cypriniformes; Cyprinidae; Danio.
REFERENCE 1  (bases 1 to 776)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL   Contact: Daniela S. Gerhard, Ph.D.
COMMENT   Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Leonard I. Zon, M.D.
          cDNA Library Preparation: Invitrogen Corp
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM14604 row: n column: 12
          High quality sequence stop: 698.
FEATURES             Location/Qualifiers
     source          1..776
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:6964621"
                     /tissue_type="embryo"
                     /lab_host="DH10B (T1-resistant)"
                     /clone_lib="NCI_CGAP_ZEmb2"
                     /note="Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI;
                     Cloned unidirectionally. Primer: Oligo dT. Average insert
                     size 2 kb. Constructed by J. Wang (Research Genetics,
                     Invitrogen Corp) from tissue donated by L. Zon (Harvard
                     University). Note: this is a NCI_CGAP Library."
ORIGIN
Alignment Scores:
Pred. No.:           3.61e-98           Length:          776
Score:               925.00             Matches:         181
Percent Similarity:  83.20%             Conservative:    27
Best Local Similarity: 72.40%           Mismatches:      42
Query Match:         63.62%             Indels:          0
DB:                                     Gaps:            0

US-09-889-733B-2 (1-285) x CD751679 (1-776)
```

```
Qy     8 SerAlaPheGlyArgAlaThrHisAlaValAlaValArgAlaLeuProLeuSerLeuCysGln 27
                :::      :::::::::   ::::::|||||:::::|||
Db    26 TCCGACTTCGGCAAGTACACGCACGCGATCATGCGCGGATGCCCGGTCTCTGAGCACC 85

Qy    28 HisAlaLeuArgSerAlaLysGlyGlyValGluValAspValAlaArgAlaArgGlnHis 47
                                           ::::|||||:::
Db    86 GCCGCCGCTGCGCTCCGATAATGTGTCCCCGGTGGATCCCCGTCGCGGTGCAGCGGGAGTTT 145

Qy    48 GlnLeuTyrValGlyValLeuGlySerLysLysGlyLeuGlnValValGluLeuProAla 67
                                       :::|||||::::
Db   146 GACCACTATGTGTCCCGTGCTCAGAGGACGGGCTCGGGCTGCAGGTGGTGATGCTGCCCGCT 205

Qy    68 AspGluSerLeuProAspCysValPheValGluArgValAspAspMetLysMetLysGlu 87
                                       :::::|||||:::
Db   206 GACGAGGAGGGTCTCCATCTGAAGAGCTTCGCAGTATGGCTGGTCCAAACCTCATCGCC 265

Qy    88 AlaLeuIleThrArgProGlyAlaProSerArgArgArgLysGluValAspMetLysGlu 107
                                               ||||::::
Db   266 GCGCTCATCACACCAGACCCCGAGACCCCGAGACCGGCCGGTCGAGACCGTGGCCATGAAGGAT 325

Qy   108 AlaLeuGluLysLeuGlnLeuAsnLeuAsnLeuValGluMetLysAspGluAsnAlaThrLeuAsp 127
                                           |||:::
Db   326 GCGCTGACTGACGCTCGGCCCTGAACATTGTGGAGATCGAATGATGAATCAGCCACGATGGAT 385

Qy   128 GlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgArgThrAsn 147
```

```
Db   386 GGAGGAGCAGCGTCCTGTTCACAGGAAAGAGTTCTTTGTGGGATTATCAAAAGAGAACCAAT 445
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   148 GlnArgGlyIleAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrValPro 167
             ||||||::::      ::::::      :::      ||||
Db   446 CAGAGAGGAGCAGAGATCCTGGCCAACACCTTAAGGATTACGCTGGTCCTATCCCT 505

Qy   168 ValAlaAspGlyLeuHisLeuLysSerSerPheCysSerMetAlaGlyProAsnLeuUIleAla 187
             :::|||||||||||                   :::::::::
Db   506 GTAGAGGAGGGGTCTCCATCTGAAGAGCTTCGCAGTATGGCTGGTCCAAACCTCATCGCC 565

Qy   188 IleGlySerGluSerArgIleAlaGlnLysAlaLeuUIleMetGlnMetSerAspHis 207
             :::::::::::::::::            :::
Db   566 ATCGGCTCCAGTGATGCCGCCACAGAGGCCCTGAAGGTAATACAGCAGATGAGTGATTGT 625

Qy   208 ArgTyrAspLysProAspAspProAspAspAlaAsnCysIleTyrLeuAsnIle 227
             :::::::::::
Db   626 AAATATGGAGAAGCTGACGTTCCTGATGATCGGGTCGACAGTGTGTACATGAACCTG 685

Qy   228 ProAsnLysGlyHisValLeuHisLeuHisArgThrProGluGluTyrProGluSerAlaLys 247
             :::::::::::
Db   686 CCGGGGAAAGGAGGAGACGTTCTGCTCACTGCACACCTGAAGTGTTTCCTGAGAGCGCCNAG 745

Qy   248 ValTyrGluLysLeuLysAspHisMetLeu 257
             :::::::
Db   746 GTGTTCGAGAGGCTGAAGAGTCACATGCTC 775
```

```
RESULT 13
BB611533
LOCUS       BB611533                 731 bp    mRNA        linear   EST 26-OCT-2001
DEFINITION  BB611533 RIKEN full-length enriched, 13 days embryo head Mus
            musculus cDNA clone 3110053015 5', mRNA sequence.
ACCESSION   BB611533
VERSION     BB611533.1  GI:16452732
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 731)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center (GSC) Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K..
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res.  10 (10), 1617-1630 (2000)
            wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
            Computer-based methods for the mouse full-length cDNA
```

Left column

```
          encyclopedia: real-time sequence clustering for construction of a
          nonredundant cDNA library. Genome Res. .11 (2), 281-289 (2001)
          Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
          Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
          Hayashizaki,Y.
          Computational Analysis of Full-Length Mouse cDNAs Compared with
          Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
          Please visit our web site (http://genome.gsc.riken.go.jp) for
          further details.
          e mouse tissues.
FEATURES             Location/Qualifiers
     source          1..731
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="311005301S"
                     /lab_host="head"
                     /tissue_type="head"
                     /dev_stage="13 days embryo"
                     /clone_lib="RIKEN full-length enriched, 13 days embryo
          head"
                     /note="Site 1: XhoI; Site 2: SstI; cDNA library was
          prepared and sequenced in Mouse Genome Encyclopedia
          Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in
          RIKEN. Division of Experimental Animal Research in Riken
          contributed to prepare mouse tissues. 1st strand cDNA was
          primed with a primer [5'
          GAGAGAGAGAGAGAGATCCAAGAGCTCTTTTTTTTTTTTTTVN 3'], cDNA was
          prepared by using trehalose thermo-activated reverse
          transcriptase and subsequently enriched for full-length by
          cap-trapper. cDNA went through one round of normalization
          to Rot = 10.0 and subtraction to Rot = 50.0. Second strand
          cDNA was prepared with the primer adapter of sequence [5'
          GAGAGAGAGAATTCTCAGTTAATTAATTAATCCCCCCCCCCCCCC 3']. cDNA
          was cleaved with XhoI and SstI."
ORIGIN

Alignment Scores:
Pred. No.:                    8.66e-95          Length:           731
Score:                        896.00            Matches:          180
Percent Similarity:           94.90%            Conservative:     6
Best Local Similarity:        91.84%            Mismatches:       9
Query Match:                  61.62%            Indels:           1
DB:                                             Gaps:             0

US-09-889-733B-2 (1-285) x BB611533 (1-731)

Qy    1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  134 ATGGCCGGCCTCGGCCACCCCTCCGCCTTCGGCCGCGCCACCCACGCCGTGGTGCGGGCT 193

Qy   21 LeuProGluSerLeuCysGlnHisAlaAlaLeuArgSerAlaLysGlyGluValAspVal 40
          |||||||||||||||||||||:::::|||||||||||||       |::::||||||||
Db  194 CCGCCCGAGTCCCTGTGCCACCACGCGGCGCTGAGGGCGTCGCAGGGCGAGGAGGTGGATTC 253

Qy   41 AlaArgAlaGluArgGlnHisHisGlnLeuTyrValGlyValAlaGlySerLysGlyLeu 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  254 GCTCGCGCCGGAGCGCCAGCACCACGAGCTCTACGTGGGCGTGGCCAGCAAGCTGGGGCTG 313

Qy   61 GlnValValGluProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
          ||:|||||::|::|||||||:|||||||||||||||||||||||||||||||||||
Db  314 CAGGTGGTGCAGCTGCCCGCCGACGAGAGCCTGCCCGACTGCGTGTTCGTGGAGGACGTG 373

Qy   81 AlaAlaValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
          ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db  374 GCCGCCGTGTGCGAGGAGACGGCCCTCATCACCCGCCCGGGCGCCCCCAGCCGCAGGAAG 433

Qy  101 GluValAspMetMetLysGlyGluAlaLeuGlyLysGlnLeuLeuAsnIleValGluLys 120
          |||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db  434 GAGGTTGACATGATGAAAGAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAA 493

Qy  121 AspGluAsnAlaThrLeuAspGlyGlyValAspValAspGlyArgGluPhePheVal 140
```

Right column

```
Db  494 GATGAAAATGCAACTTTGGATGGTGGGACGTCCCATTCACAGGCAGGANGAATTTTGTG 553
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  141 GlyLeuSerLysArgThrAsnGlnArgGlyValAlaGluIleLeuAlaAspThrPheLysAsp 160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  554 GGCCTTTCCAAAAGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 613

Qy  161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSer-Me 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  614 TACGCAGTCTTCACAGATCCTGTGGCCGATTCTTTGCATTTAAAGAGATTCTGCAGCCAT 673

Qy  180 tAlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGln 195
          |||||||||||||||||||||||||||||::::|||||||||||||
Db  674 GGCCGGGACCCAACCTGATTGCATTAGGTCCAGCGAATCTGCACAG 719

RESULT 14
CK019367
LOCUS       AGENCOURT_16541797 NIH_ZGC_7  897 bp      mRNA     linear     EST 26-NOV-2003
DEFINITION  5', mRNA sequence.
ACCESSION   CK019367
VERSION     CK019367.1  GI:3B545291
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 897)
            NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished (1999)
JOURNAL     Contact: Daniela S. Gerhard, Ph.D.
COMMENT     Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Len Zon, Harvard
            cDNA Library Preparation: Open Biosystems
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14786 row: g   column: 11
            High quality sequence stop: 609.
FEATURES             Location/Qualifiers
     source          1..897
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:7037029"
                     /tissue_type="whole body"
                     /lab_host="DH10B"
                     /clone_lib="NIH_ZGC_7"
                     /note="Vector: pExpress1; Site 1: EcoRV;
                     Site 1: NotI; Site 2: EcoRV;
                     /note="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
                     Bulk tissue was collected from a whole adult individual
                     from the Tuebingen strain.\1st strand cDNA was primed with
                     a Not I - oligo(dT) primer, double-stranded cDNA was
                     cloned into the Not I and EcoRV sites of pExpress-1.
                     Library was size-selected for >1 kb fragments and
                     normalized. A non-normalized version of this library is
                     also available (NIH_ZGC_10). Library was constructed by
                     Open Biosystems (Huntsville, AL)."
ORIGIN

Alignment Scores:
Pred. No.:                    3.15e-94          Length:            897
Score:                        892.50            Matches:           185
Percent Similarity:           80.51%            Conservative:      34
Best Local Similarity:        68.01%            Mismatches:        49
Query Match:                  61.38%            Indels:            6
DB:                                             Gaps:              3
```

FEATURES   Location/Qualifiers
     source   1..923
                 /organism="Mus musculus"
                 /mol_type="mRNA"
                 /strain="FVB/N"
                 /db_xref="taxon:10090"
                 /clone="IMAGE:5134831"
                 /lab_host="DH10B (T1 phage-resistant)"
                 /clone_lib="NCI_CGAP_Li9"
                 /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                 Average insert size 1.9 kb. Constructed by Life
                 Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:
Pred. No.:
Score: 892.00   Length: 923
Percent Similarity: 84.96%   Matches: 215
Best Local Similarity: 80.83%   Conservative: 11
Query Match: 61.35%   Mismatches: 23
DB: 12   Indels: 19
                         3.78e-94   Gaps: 4

US-09-889-733B-2 (1-285) x BI332932 (1-923)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaAlaValValArgAla   20
        ATGGCCGGGCCTCGGCCACCCCTCCGGCTTCGGCCGCGGCCACCCACGCGTGGTGCGGGCT
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 149 ATGGCCGGGCCTCGGCCACCCCTCCGGCTTCGGCCGCGGCCACCCACGCGTGGTGCGGGCT 208

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSer-AlaLysGlyGluValAspVa   40
        CCGCCCGAGTCCCTGTGCCGCCACGCGCTGAGTCGCCGC-CGAGGGCCGGAGGAGGTGGATTT
        |||||||||||||||||||||||||::::::|||||||||||||||||||||||||||
Db 209 CCGCCCGAGTCCCTGTGCCGCCACGCGCTGAGTCGCCGCCGAGGGCCGGAGGAGGTGGATTT 268

Qy   40 1AlaArgAlaGlnHisGlnLeuTyrValGlyValLeuGlySerLysLysGlyLe   60
        CGCTCGGCGCCGCAGCACCAGCTCTACGTCGGGGTGCTGGGCAGCAAGAGCTCGGGGCT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 269 CGCTCGGCGCCGCAGCACCAGCTCTACGTCGGGGTGCTGGGCAGCAAGAGCTCGGGGCT   328

Qy   60 uGlnValValGlu-LeuProAlaAspGluSerLeu-ProAspCysValPheValGluAsp   79
        GCAGGTGGTGGAG-CTGCCTGCCGACGAGAGCCTC-CCCGACTGCGTGTTCGTGGAGGAC
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 329 GCAGGTGGTGGAGCTGCCTGCCGACGAGAGCCTCCCCGACTGCGTGTTCGTGGAGGAC   388

Qy   80 Val-AlaValValCysGluGluThrAlaLeuIleThrArg--------ProGlyAl   95
        GTT-GCCGTCGTGTGCGAGGAGACGGCCCTCATCACCCGCTGCCTATGC
        ||||||||||||||||||||||||||||||||||||||::::|||||
Db 389 GTTGCCGTCGTGTGCGAGGAGACGGCCCTCATCACCCCGCTGCCTATGC   448

Qy   95 aProSerArgArgLysGluValValAspMetMetLysGluLysLeuGlnLeuAs   115
        ------GAGGTTGACATGATGAAAGAGAGCTTTGGAAAAACTTCAGCTCAA
        ::::::|||||||||||||||||||||||||||||||||||||||||
Db 449 GAGGAGG------GAGGTTGACATGATGAAAGAGAGCTTTGGAAAAACTTCAGCTCAA   502

Qy 115 nIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGly-LeuPheThrG   135
        CATGTAGAGATGAAAGATGAAAATGCAACTTTGGATGTGGGGGACGTTCCTATTCACAG
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 503 CATGTAGAGATGAAAGATGAAAATGCAACTTTGGATGTGGGGGACGTTCCTATTCACAG   562

Qy 135 lyArgGluPheValGlyLeuSerLysArgThrAsnGlnArgGly-AlaGluIleLeuLeu   154
        GCAGAGAATTTTTGTGGGCCTTTCCAAAAGAACAAATCAACAAGGAGGTTGCTGAAATCTTG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 563 GCAGAGAATTTTTGTGGGCCTTTCCAAAAGAACAAATCAACAAGGAGGTTGCTGAAATCTTG 622

Qy 155 AlaAspThrPheLysAspTyrAlaValSerThrValProValAlaAspGlyLeuHisLeu 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db    623  GCTGATACTTTAAAGGACTACGCAGTCTCTACAGTCCCTGTGGCCGATTCTTTGCATTTA  682

Qy    175  LysSerPheCysSerMet-AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAl  194
                |||||||||||||||||          |||||||||||||||||||||||
Db    683  AAGAGTTTCTGCAGCATGCCCGGACCCAACCTGATTGCAATAGGGTCCAGCGAATCTTG  742

Qy    194  aGln---LysAlaLeuLysIleMetGlnGln-MetSerAspHisArgTyrAspLysLeuT  213
              :::|           || :::|||||||||||||||||||||||||||
Db    743  CACAGGAAGGGCCTCAAAGATCATTGCACAGATTGAGTGACCATCGGTATGACAAGCTCA  802

Qy    213  hrVal-ProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHis  232
              |||| |||          |||||||  ||||||||      ||||
Db    803  CTGTTACCCGGAGACACATTGCCGGCAACTGCATAATATTAAATATCCCAGCAAAGGCA--  859

Qy    233  ValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLeu  252
              ||||||   |||          |||||||||:::  ||||||   ||||
Db    860  TGTCTTCTGGACCGAAC-CCAGAAGAGTACCCAAAAAGCGCAAGGGC-TATGAGAAACTTA  917

Search completed: June  8, 2004,  01:05:11
Job time : 2536 secs
```

```
              GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.
```

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2004, 23:13:51 ; Search time 89 Seconds
               (without alignments)
               1777.090 Million cell updates/sec

Title: US-09-889-733B-2
Perfect score: 1454
Sequence: 1 MAGLGHPSAFGRATHAVVRA........ERKVDGLLTCCSVLINKKVDS 285

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext 0.5
               Ygapop 10.0 , Ygapext 0.5
               Fgapop 6.0 , Fgapext 7.0
               Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

```
Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09889733/runat_07062004_083529_8000/app_query.fasta_1.455
-DB=Issued_Patents_NA__QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09889733_@CGN_1_1.105_@runat_07062004_083529_8000 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY-NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
  1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
  2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
  3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
  4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
  5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
  6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	645.5	44.4	1397	4	US-09-023-655-37	Sequence 37, Appl
2	301	20.7	2598	4	US-09-252-991A-2768	Sequence 2768, Ap
3	301	20.7	2790	4	US-09-252-991A-2864	Sequence 2864, Ap
4	233	16.0	239	4	US-09-702-705-404	Sequence 404, App
5	233	16.0	239	4	US-09-736-457-404	Sequence 404, App
6	233	16.0	239	4	US-09-614-124B-404	Sequence 404, App
7	233	16.0	239	4	US-09-671-325-404	Sequence 404, App
8	233	16.0	239	4	US-09-589-184-404	Sequence 404, App
9	198	13.6	526	4	US-09-621-976-12670	Sequence 12670, A
10	146.5	10.1	513	4	US-09-621-976-370	Sequence 370, App
11	122.5	8.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
12	122.5	8.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
13	113.5	7.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
14	113.5	7.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
15	102	7.0	1390	3	US-08-605-150A-7	Sequence 7, Appli
16	101.5	7.0	1354	3	US-08-605-150A-1	Sequence 1, Appli
17	97.5	6.7	1263	4	US-09-134-001C-1590	Sequence 1590, Ap
18	97.5	6.7	1440	4	US-09-252-991A-4726	Sequence 4726, Ap
19	97.5	6.7	2780	4	US-09-620-312D-358	Sequence 358, App
20	97.5	6.7	3060	4	US-09-252-991A-4686	Sequence 4686, Ap
21	97.5	6.7	3456	4	US-09-252-991A-4757	Sequence 4757, Ap
22	93	6.4	1708	4	US-09-919-172-30	Sequence 30, Appl
23	93	6.4	1709	2	US-09-010-398-2	Sequence 2, Appli
24	93	6.4	1709	3	US-09-166-260-2	Sequence 2, Appli
25	93	6.4	1709	3	US-09-153-804-11	Sequence 11, Appl
26	92.5	6.4	1490	3	US-08-605-150A-5	Sequence 5, Appli
27	90.5	6.2	1464	3	US-08-605-150A-3	Sequence 3, Appli
28	90.5	6.2	2807	4	US-09-453-702B-51	Sequence 51, Appl
29	90	6.2	1068	4	US-09-328-352-434	Sequence 434, App
30	89.5	6.2	1545	2	US-08-628-039-1	Sequence 1, Appli
31	89.5	6.2	1545	2	US-08-628-039-5	Sequence 5, Appli
32	89.5	6.2	1545	2	US-08-628-039-7	Sequence 7, Appli
33	89.5	6.2	1545	2	US-08-628-039-8	Sequence 8, Appli
34	89.5	6.2	1545	2	US-08-673-388-1	Sequence 1, Appli
35	89.5	6.2	1545	2	US-08-673-388-5	Sequence 5, Appli
36	89.5	6.2	1545	2	US-08-673-388-7	Sequence 7, Appli
37	89.5	6.2	1545	2	US-08-673-388-8	Sequence 8, Appli
38	89.5	6.2	1545	2	US-08-614-877-1	Sequence 1, Appli
39	89.5	6.2	1545	2	US-08-614-877-5	Sequence 5, Appli
40	89.5	6.2	1545	2	US-08-614-877-7	Sequence 7, Appli
41	89.5	6.2	1545	3	US-08-614-877-8	Sequence 8, Appli
42	89.5	6.2	1545	3	US-08-912-205-1	Sequence 1, Appli
43	89.5	6.2	1545	3	US-08-912-205-5	Sequence 5, Appli
44	89.5	6.2	1545	3	US-08-912-205-7	Sequence 7, Appli
45	89.5	6.2	1545	3	US-08-912-205-8	Sequence 8, Appli

ALIGNMENTS

```
RESULT 1
US-09-023-655-37
; Sequence 37, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
;  APPLICANT: Cocks, Benjamin G.
;  APPLICANT: Susan G. Stuart
;  APPLICANT: Jeffrey J. Seilhamer
;  TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;  TITLE OF INVENTION: EXPRESSION
;  NUMBER OF SEQUENCES: 1508
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/023,655
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
;  TELECOMMUNICATION INFORMATION:
```

```
      TELEPHONE:    (650) 855-0555
      TELEFAX:      (650) 845-4166
  INFORMATION FOR SEQ ID NO: 37:
     SEQUENCE CHARACTERISTICS:
        LENGTH:  1397 base pairs
        TYPE:  nucleic acid
        STRANDEDNESS:  single
        TOPOLOGY:  linear
     IMMEDIATE SOURCE:
        LIBRARY:  HMC1NOT01
        CLONE:    003437
US-09-023-655-37

Alignment Scores:
Pred. No.:                              Length:           1397
Score:               1.68e-72           Matches:          141
Percent Similarity:  645.50             Conservative:     46
Best Local Similarity: 68.75%           Mismatches:       78
Query Match:         51.84%             Indels:           7
DB:                  44.39%             Gaps:             3

US-09-889-733B-2 (1-285) x US-09-023-655-37 (1-1397)

Qy    3  GlyLeuGlyHisPro---SerAlaPheGlyAlaArgAlaThrHisAlaValValArgAlaLeu  21
                  ::||:||| |||                    :::|||||       :::
Db   332  GGGATGGGGACCCCGGGAGGGGGGCTGGCCGCTGCTCCCATGCCAGCCCTGACCCCGGAGTC  391

Qy   22  ProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspValAla    41
                 :::||||||||                   ::::|||::::
Db   392  CCAGAGAGCCTGGCGTCGGGGGAAGGTGCGGGGGGCTGGCCTTCCCGCTCTGGATCTGGCC  451

Qy   42  ArgAlaGluArgGlnHisAlaLeu                   SerLysLeuGlyLeuGln    61
             ::|:||:||||                               :::|||||||
Db   452  AAAGCTCAAAGGGAGCACGGGGTGCTGGGAGGTAAACTGACTGGGGCTACAG            511

Qy   62  ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla    81
                :::                    :::|||:::|||||||
Db   512  CTGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGGACCGGCTCGTTGCGCACACGGCC  571

Qy   82  ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgArgGlu   101
            ||||||||                              :::|||:::
Db   572  GTGATCCAAGGGACACGGGGCCTAATCACGGGGCCCTGGAGCCCCGCTCGTAGGCCAGAG   631

Qy  102  ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp   121
              :::|:|::||||                   ::|::|||||
Db   632  GTCGATGGAGTGCCCAAAGCCCTGCAAGACCCTGGGGCTTCCGAATTGTGAAATAGGAGAC  691

Qy  122  GluAsnAlaThrLeuAspGlyGlyAspValProValPheGlyArgGlnPhePheValGly   141
              :::|||                    :::||||||::
Db  692  GAGAACGCGACGCTGGATGGCACTGGACGTTCTCTTCACCGGCCGGGAGTTTTTCGTAGGC  751

Qy  142  LeuSerLysArgThrAsnGlnArgGlyAla--GluIleLeuAlaAlaAspThrPheLysAspTy  161
              ::||||                         :::||||||:::|:::
Db  752  CTCTCCAAATGGACCAATCACCGAGGCGCTGGAGATCGTGGCGGCAGACGTTCCGGGACTT  811

Qy  161  rAlaValSerThrValProValAlaAlaAspGlyLeuHisLeuLysSerPheCysSerMetAl  181
             :::|::|||                    :::||:::
Db  812  CGCCGTTCCACTGTCCCAGTGTGGCGGTTCCTCGGGTCCCTCTCGCCGGCATGGG        871

Qy  181  aGlyProAsnLeuIleAlaIleGlyTyrAspLysSerGluSerAlaGlnAlaLeuLysIleMe  201
             ::::|||                    :::||||||
Db  872  GGGACCTTCGCGCACTGTTGTGCAGGCAGCAGCCGCTGCCCAAAAGGCTGTCCGGGCAAT  931

Qy  201  tGlnIleMetSerAspHisArgTyrAspLysLeuThr--ValProAspAlaAlaAlaA    221
              ::::|||                    ::::||||:::|
Db  932  GGCAGTGCTGACAGATCACCCCATAATGCCCTCCCTGACCCCCCCAGATGACGCAGCTGCTG  991

Qy  221  snCysIleTyrLeu------AsnIleProAsnLysGlyHisValLeuLeuHisArgThrP  239
              :::|:::                  ::::|||||:::
Db  992  ACTGTCTCTTTCTTTCGTCCTGGGTTGCCTGGTGCCCCTTTCCCTCCTCCACCGTGGAG  1051

Qy  239  roGluGlu-------TyrProGluSerAlaLysValTyrGluLysLysAspHisMetL  257
               :::                  ::  ||:::|||||:::
Db  ...  CACCTGAGACCCGGGCTGCCTACCTGAACACAACACCTGCTGGCCGCCGGCGAGTC  ...
```

```
Db  1052  GTGGGGATCTGTGCNCCCAAACAGCCAGGAGGCACTGCGAGAAGCTTCTGATGTCACCC  1111

Qy   257  euIleProValSerMetSerGluLeuGluLys  267
              |:::|||||                :::
Db  1112  TGGTACCTGTGTCCTGCTGCTCAGAACTGAGAAG  1143
```

RESULT 2
US-09-252-991A-2768
; Sequence 2768, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;   APPLICANT: Marc J. Rubenfield et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;   TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 107196.136
;   CURRENT APPLICATION NUMBER: US/09/252,991A
;   CURRENT FILING DATE: 1999-02-18
;   PRIOR APPLICATION NUMBER: US 60/074,788
;   PRIOR FILING DATE: 1998-02-18
;   PRIOR APPLICATION NUMBER: US 60/094,190
;   PRIOR FILING DATE: 1998-07-27
;   NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2768
;   LENGTH: 2598
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2768

```
Alignment Scores:
Pred. No.:                              Length:           2598
Score:               5.56e-28           Matches:          82
Percent Similarity:  301.00             Conservative:     47
Best Local Similarity: 48.50%           Mismatches:       117
Query Match:         30.83%             Indels:           20
DB:                  20.70%             Gaps:             8

US-09-889-733B-2 (1-285) x US-09-252-991A-2768 (1-2598)

Qy   15  HisAlaValValArgAlaLeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLys   34
             ::|||                    :::|||
Db  320  CACATCATCGCTCGCACGCCCCCGGCGCTGGTCGACGGGCCTGACCTCCAGCCACCTC   379

Qy   35  GlyGluGluValAspValAlaArgAlaGluArgGlnHisGlyLeuTyrValGlyValLeu   54
             ||||||                    :::||||||
Db  380  GGCAAGCCG---GACTACGCCAAGGCCTGGAGCAGCAGCAACGCCTACATCCGCGCCCTG   436

Qy   55  GlySerLysLeuGlyLeuGlnValValGluLeuProAlaAspGluSerLeuProAspCys   74
             :::|||                    ::|||
Db  437  ---CAGACCTGCGACGTGACATCACCCCGTCTCCGCGGGACGAACGCTTCCCGACTCG   493

Qy   75  ValPheValGluAspValAlaValValCysGluGluThrAlaLeuIleThrArgProGly   94
             ||||||                    :::|||
Db  494  GTGTTCGTCGAGGACCCGGTGCTCTGCACCTGCGCCATCATCACCCGGCCCCCCGGC   553

Qy   95  AlaProSerArgArgLysGluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeu  114
             ::||||                    :::|||
Db  554  GCCGAATACGCGGCCGCGGCGGGAGACCGAGATCATCGAGGAAACCGTGCAGCGCTTCATCCG   613

Qy  115  AsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValProValPheThr  134
             ::|||                    :::|||
Db  614  GGCAAGGTCGAGGCGCATCGAGGCACCCGGCACCGTGGAAGCCGTGAGCATCATGATGGTC   673

Qy  135  GlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeu  154
             ::|||                    :::|||
Db  674  GGCGACCACTTCTACATCGGCGAATCGGCAGAATCGGCCACCAACGGCCCGAGGGGCCCAGATG   733

Qy  155  AlaAspThrPheLysAspTyrAlaVal-----SerThrValProValAlaAspGlyLeu  172
             ::|||                    :::
Db  734  ATCCGGATCCTGGAGAAACATGGCCTCACGGGCTTCGGTGGTGCCCTGAAAAGGTCCTG   793

Qy  173  HisLeuLysSer---PheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSer  191
             :::|||
Db  794  CACCTGAGACCCGGGCTGCCTACCTGAACACAACACCTGCTGGCCGCCGGCGAGTC  853
```

Right column

Qy 115 AsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThr 134
                                    ::: |||
Db 2329 GGCAGGTCGAGCGCCGATCGAGGCACCCGGCACGGTGGAAGCCGGCGACATCATGATGGTC 2388

Qy 135 GlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyIleAlaGluIleLeu 154
                |||:   :::::::|||:||||
Db 2389 GGCGACCACTTCTACATCGGGAATCGGCCCGCACCACGCGGAGGGGCCCGGCAGATG 2448

Qy 155 AlaAspThrPheLysAspTyrAlaVal------SerThrValProValAlaAspGlyLeu 172
                ||||| :::
Db 2449 ATCGGCATCCTGGAGAAACATGGCCTCAGCGGCTCGGTGTGGCCTGGAAAAGGTCCTG 2508

Qy 173 HisLeuLysSer---PheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSer 191
                |||||:::
Db 2509 CACCTGAAGACCGGGCTCGCCTACCTGGAACACAACCTTGCTGCCGCCGGCGAGTTC 2568

Qy 192 GluSerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisArgTyrAspLys 211
                                    :::   :::
Db 2569 GTCAGC----------------------AAGCCGGAGTTCCAGGACTTCAACATCATCGAG 2607

Qy 212 LeuThrValProAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGly 231
                ::: ||||||:::::::|||:
Db 2608 ATCCCCGAAGAGGAGTCCTACGCCGCCAACTGCATCTGGGTCAAC-----------GAA 2655

Qy 232 HisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLys 251
                ||:::::   :::
Db 2656 AGGGTGATCATG------------CCCGCGGCTATCCCCGGACCCGGAGAAGATCGCCGC 2706

Qy 252 LeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeu 271
                |||
Db 2707 CTC---GGCTACCGGGTGATCGAGGTGGACACCTCCGAATATCGCAAGATCGACGCGGGC 2763

Qy 272 LeuThrCysCysSerVal 277
                :::::|||:
Db 2764 GTCAGTTGCATGTCGCTG 2781

RESULT 4
US-09-702-705-404/c
; Sequence 404, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   APPLICANT: Fan, Liqun
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C14
;   CURRENT APPLICATION NUMBER: US/09/702,705
;   CURRENT FILING DATE: 2000-10-30
;   NUMBER OF SEQ ID NOS: 1833
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
;   LENGTH: 239
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-702-705-404

Alignment Scores:
Pred. No.:
Score: 5.85e-21
                         233.00
Percent Similarity: 78.48%   Length: 239
Best Local Similarity: 62.03%   Matches: 49
Query Match: 16.02%   Conservative: 13
DB: Mismatches: 17
                                         Indels: 0
                                         Gaps: 0

US-09-889-733B-2 (1-285) x US-09-702-705-404 (1-239)

Left column

Qy 192 GluSerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisArgTyrAspLys 211
                                    ::: |||
Db 854 GTCAGC----------------------AAGCCGGAGTTCCAGGACTTCAACATCATCGAG 892

Qy 212 LeuThrValProAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGly 231
                |||:   ::: ||||:||||||
Db 893 ATCCCCGAAGAGGAGTCCTACGCCGCCAACTGCATCTGGGTCAAC-----------GAA 940

Qy 232 HisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLys 251
                ||:::||| :::   :::
Db 941 AGGGTGATCATG------------CCCGCGGCTATCCCCGGACCCGGAGAAGATCGCCGC 991

Qy 252 LeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeu 271
                ::: |||:::||:|||
Db 992 CTC---GGCTACCGGGTGATCGAGGTGGACACCTCCGAATATCGCAAGATCGACGCGGGC 1048

Qy 272 LeuThrCysCysSerVal 277
                :::::|||:|||:|
Db 1049 GTCAGTTGCATGTCGCTG 1066

RESULT 3
US-09-252-991A-2864
; Sequence 2864, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;   APPLICANT: Marc J. Rubenfield et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;   TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 107196.136
;   CURRENT APPLICATION NUMBER: US/09/252,991A
;   CURRENT FILING DATE: 1999-02-18
;   PRIOR APPLICATION NUMBER: US 60/074,788
;   PRIOR FILING DATE: 1998-02-18
;   PRIOR APPLICATION NUMBER: US 60/094,190
;   PRIOR FILING DATE: 1998-07-27
;   NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2864
;   LENGTH: 2790
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2864

Alignment Scores:
Pred. No.:
Score: 6.24e-28
                         301.00
Percent Similarity: 48.50%   Length: 2790
Best Local Similarity: 30.83%   Matches: 82
Query Match: 20.70%   Conservative: 47
DB: Mismatches: 117
                                         Indels: 20
                                         Gaps: 8

US-09-889-733B-2 (1-285) x US-09-252-991A-2864 (1-2790)

Qy   15 HisAlaValValArgAlaLeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLys 34
                            :::   ::: |||||
Db 2035 CACATCATCGCTCGCAGCCCGCCGCCCGCCGGTCGACGGCCTGACCTCCAGCCACCTC 2094

Qy   35 GlyGluGlyGluAspValAlaArgAlaGluAlaArgGlnHisGlnHisLeuTyrValGlyIleValLeu 54
                ||||:::   :::||:::
Db 2095 GGCAAGCCG---GACTACGCCAAGGCCCTGGAGCAGCAGCAACGGCCTACATCCGGGCCTTG 2151

Qy   55 GlySerLysLeuGlyLeuValValGlnLeuValGluLeuProAlaAspGluSerLeuProAspCys 74
                ::: :::
Db 2152 ---CAGACCTGCGACGTGGACATCACCCTGCTGCCGCCGGCGACGAACGCTTCCCCGACTCG 2208

Qy   75 ValPheValGluAspValAlaValAlaValCysGluGluThrAlaGluIleThrArgProGly 94
                :::   :::   :::
Db 2209 GTGTTCGTGGAGGACCCCGGTGCTTCGCACCTGCGTCGCCATCATCACCCGCCCGGC 2268

Qy   95 AlaProSerArgArgLysGluGluValAlaAspMetGlyMetLysGluAlaLeuGlyLeuLysGluLysLeuGlnLeu 114
                ||||||||:::   :::
Db 2269 GCCGAATCGGCGCGCGGCGAGACCGAGATCATCGAGGAAACCGTGCAGCGCGCGCTTCTCATCCG 2328

Qy   62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
          :::::||||||||| ::: ||| |||
Db   239 CTGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCCGTGCTTGGCGACACGGCC 180

Qy   82 ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
          ||||:   :::||||:::|||:::|||||||
Db   179 GTGATCCAAGGGACACGGCCCTAATCACGCGGCCCTCGGAGCCCGCCTCGTAGGCCAGAG 120

Qy   102 ValAspMetMetLysGluAlaLeuGluLysLysGlnLeuAsnIleValGluMetLysAsp 121
          :::::|||||||||:::
Db   119 GTCGATGGAGTCCGCAAAGCCCTGCAAGACCTGGGGCTCGAATTGTGGAAATAGGAGAC 60

Qy   122 GluAsnAlaThrLeuAspGlyGlyValLeuPheThrGlyArgGluPhePheVal 140
          |||||||||||||||||||||||||||||||||||||||||||
Db   59 GAGAACGCGACCGCTGGATGGCCACTGACGTTCTCTTCACCGGCCGGGAGTTTTTCGTA 3

RESULT 5
US-09-736-457-404/c
; Sequence 404, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-404

Alignment Scores:
Pred. No.:   5.85e-21   Length: 239
Score: 233.00   Matches: 49
Percent Similarity: 78.48%   Conservative: 13
Best Local Similarity: 62.03%   Mismatches: 17
Query Match: 16.02%   Indels: 0
DB: Gaps: 0

US-09-889-733B-2 (1-285) x US-09-736-457-404 (1-239)

Qy   62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
          :::::||||||||| ::: ||| |||
Db   239 CTGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCCGTGCTTGGCGACACGGCC 180

Qy   82 ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
          ||||:   :::||||:::|||:::|||||||
Db   179 GTGATCCAAGGGACACGGCCCTAATCACGCGGCCCTCGGAGCCCGCCTCGTAGGCCAGAG 120

Qy   102 ValAspMetMetLysGluAlaLeuGluLysLysGlnLeuAsnIleValGluMetLysAsp 121
          :::::|||||||||:::
Db   119 GTCGATGGAGTCCGCAAAGCCCTGCAAGACCTGGGGCTCGAATTGTGGAAATAGGAGAC 60

Qy   122 GluAsnAlaThrLeuAspGlyGlyValLeuPheThrGlyArgGluPhePheVal 140
          |||||||||||||||||||||||||||||||||||||||||||
Db   59 GAGAACGCGACCGCTGGATGGCCACTGACGTTCTCTTCACCGGCCGGGAGTTTTTCGTA 3

RESULT 6
US-09-614-124B-404/c
;¡ Sequence 404, Application US/09614124B

; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-404

Alignment Scores:
Pred. No.:   5.85e-21   Length: 239
Score: 233.00   Matches: 49
Percent Similarity: 78.48%   Conservative: 13
Best Local Similarity: 62.03%   Mismatches: 17
Query Match: 16.02%   Indels: 0
DB: Gaps: 0

US-09-889-733B-2 (1-285) x US-09-614-124B-404 (1-239)

Qy   62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
          :::::||||||||| ::: ||| |||
Db   239 CTGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCCGTGCTTGGCGACACGGCC 180

Qy   82 ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
          ||||:   :::||||:::|||:::|||||||
Db   179 GTGATCCAAGGGACACGGCCCTAATCACGCGGCCCTCGGAGCCCGCCTCGTAGGCCAGAG 120

Qy   102 ValAspMetMetLysGluAlaLeuGluLeuLysGlnLeuAsnIleValGluMetLysAsp 121
          :::::|||||||||:::
Db   119 GTCGATGGAGTCCGCAAAGCCCTGCAAGACCTGGGGCTCGAATTGTGGAAATAGGAGAC 60

Qy   122 GluAsnAlaThrLeuAspGlyGlyValLeuPheThrGlyArgGluPhePheVal 140
          |||||||||||||||||||||||||||||||||||||||||||
Db   59 GAGAACGCGACCGCTGGATGGCCACTGACGTTCTCTTCACCGGCCGGGAGTTTTTCGTA 3

RESULT 7
US-09-671-325-404/c
; Sequence 404, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA

```
; ORGANISM: Homo sapien
US-09-671-325-404

Alignment Scores:
Pred. No.:                  5.85e-21          Length:              239
Score:                      233.00            Matches:             49
Percent Similarity:         78.48%            Conservative:        13
Best Local Similarity:      62.03%            Mismatches:          17
Query Match:                16.02%            Indels:              0
DB:                                           Gaps:                0

US-09-889-733B-2 (1-285) x US-09-671-325-404 (1-239)

Qy    62  ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValAspValAla  81
          :::|||||||||||||   |||||||||||||||||||       |||  ||||||
Db   239  CTGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCCGTCCTTGGCGACACGGCC 180

Qy    82  ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
          ||||:::    ::||||||||||||||||||||||||:::|||||||||||||||||
Db   179  GTGATCCAAGGGACACCGGCCCTAATCACGCGGCCCGCTTGGACGTCGTAGGCCAGAG 120

Qy   102  ValAspMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGlnMetLysAsp 121
          ||||||||||||||||||||||::::|||||||||||||||||::::  |||
Db   119  GTCGATGGAGTCCGCAAGACCCTGGGGCTCCGAATTGTGGAAATAGGAGAC  60

Qy   122  GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyValArgGluPhePheVal 140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    59  GAGAACGCGACGCTGGATGGCACTGACGTTCTTCACCGGCCGGGAGTTTTTCGTA  3

RESULT 8
US-09-589-184-404/c
; Sequence 404, Application US/09589184
; Patent No. 6668447
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C8
;   CURRENT APPLICATION NUMBER: US/09/589,184
;   CURRENT FILING DATE: 2000-06-05
;   NUMBER OF SEQ ID NOS: 827
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
;   LENGTH: 239
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-589-184-404

Alignment Scores:
Pred. No.:                  5.85e-21          Length:              239
Score:                      233.00            Matches:             49
Percent Similarity:         78.48%            Conservative:        13
Best Local Similarity:      62.03%            Mismatches:          17
Query Match:                16.02%            Indels:              0
DB:                                           Gaps:                0

US-09-889-733B-2 (1-285) x US-09-589-184-404 (1-239)

Qy    62  ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValAspValAla  81
          :::|||||||||||||   |||       |||  ||||||
Db   239  CTGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCCGTCCTTGGCGACACGGCC 180

Qy    82  ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
          ||||:::    ::||||||||||||||:::|||||||||||||||||
Db   179  GTGATCCAAGGGACACCGGCCCTAATCACGCGGCCCGCTTGGAGCCCGTCGTAGGCCAGAG 120
```

```
Qy   102  ValAspMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGlnMetLysAsp 121
          |||||||||||||||||||||::::|||||||||||||||||::::  |||
Db   119  GTCGATGGAGTCCGCAAAGCCCTGGGGCTCCGAATTGTGGAAATAGGAGAC  60

Qy   122  GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyValArgGluPhePheVal 140
          ||||||||||||||||||||||||||||||||||||||||||||
Db    59  GAGAACGCGACGCTGGATGGCACTGACGTTCTTCACCGGCCGGGAGTTTTTCGTA  3

RESULT 9
US-09-621-976-12670
; Sequence 12670, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Jobert, S.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;   FILE REFERENCE: GENSET.054PR2
;   CURRENT APPLICATION NUMBER: US/09/621,976
;   CURRENT FILING DATE: 2000-07-21
;   NUMBER OF SEQ ID NOS: 19335
;   SOFTWARE: Patent.pm
; SEQ ID NO 12670
;   LENGTH: 526
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-12670

Alignment Scores:
Pred. No.:                  6.32e-16          Length:              526
Score:                      198.00            Matches:             45
Percent Similarity:         73.42%            Conservative:        13
Best Local Similarity:      56.96%            Mismatches:          17
Query Match:                13.62%            Indels:              4
DB:                                           Gaps:                1

US-09-889-733B-2 (1-285) x US-09-621-976-12670 (1-526)

Qy    52  GlyValValGlySerLys---------LeuGlyLeuGlnValValGluLeuProAla  67
          ||||||||||||                 |||||||||||||||||
Db   288  GGGGTGCTGGGAGGTAAACTGAGGCAAGACTGGGGCTACAGCTGCTAGAACTGCCACCT 347

Qy    68  AspGluSerLeuProAspCysValPheValGluAspValAlaValAlaValCysGluGluThr  87
          :::||||||||||||||||||        :::   |||
Db   348  GAGGAGTCATTGCCGCTGGGACCCGTCCTTGGCGACACGGCCGTGATCCAAGGGACACG 407

Qy    88  AlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetLysGlu 107
          ||||||||||||||||||     :::||||||||::::::::
Db   408  GCCCTAATCACGCGGCCCGCTGGAGCCCGTAGGCCAGAGGTCGATGATGGAGTCCGCAA 467

Qy   108  AlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu 126
          |||||||||||||||::::|||::::   ||||||||||||||
Db   468  GCCCTGCAAGACCTGGGGCTCCGAATTGTGGAAATAGGAGACGAGAACGCGACGCTG 524
```

```
Qy   102  ValAspMetLysGluAlaLeuAlaLeuGlnLeuAsnIleValGluMetLysAsp 121
          ||||||||||||||||::::||||||||||||||::::|||
Db   119  GTCGATGGAGTCCGCAAAGCCCTGGGGCTCCGAAGACCTGGGGCTCCGAATTGTGGAAATAGGAGAC  60

Qy   122  GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyValArgGluPhePheVal 140
          |||||||||||||||||||||::::|||||||||||||||||||::::  |||
Db    59  GAGAACGCGACGCTGGATGGCACTGACGTTCTTCACCGGCCGGGAGTTTTTCGTA  3

RESULT 10
US-09-621-976-370
; Sequence 370, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Jobert, S.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;   FILE REFERENCE: GENSET.054PR2
;   CURRENT APPLICATION NUMBER: US/09/621,976
;   CURRENT FILING DATE: 2000-07-21
;   NUMBER OF SEQ ID NOS: 19335
;   SOFTWARE: Patent.pm
; SEQ ID NO 370
;   LENGTH: 513
;   TYPE: DNA
;   ORGANISM: Homo sapiens
; FEATURE:
```

; NAME/KEY: CDS
; LOCATION: 277..513
; NAME/KEY: sig_peptide
; LOCATION: 277..399
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 3.5999999463257
; OTHER INFORMATION: seq AGAGLPALDLAKA/QR
US-09-621-976-370

Alignment Scores:
Pred. No.:   2.36e-09   Length: 513
Score: 146.50   Matches: 31
Percent Similarity: 64.79%   Conservative: 15
Best Local Similarity: 43.66%   Mismatches: 24
Query Match: 10.08%   Indels: 1
DB: 4   Gaps: 1

US-09-889-733B-2 (1-285) x US-09-621-976-370 (1-513)

Qy   3 GlyLeuGlyHisPro---SerAlaPheGlyArgAlaThrHisAlaValAlaValArgAlaLeu 21
          ||||| :   ::||||||||||::||::||:::
Db 274 GGGATGGGGACCGCCGGGGGAGGGGCTGGGCCGCTGCCCATGCCCTGATCCGGGGAGTC 333

Qy   22 ProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspValAla 41
          ||| ||| || :::||||:::|||
Db 334 CCAGAGAGCCTGGCGTCGGGGGAAGGTGCGGGGGTGCGGGGCCTTGGGCCTCTGGATCTGGCC 393

Qy   42 ArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysGlyLeuGluGln 61
          :::|||||||||||| || ||| ||:
Db 394 AAAGCTCAAAGGGAGCACGGGGTGCTGGGAGGTAAACTGAGGCAACGACTGGGGCTACAG 453

Qy   62 ValValGluLeuProAlaAspAspGluSerLeuPro 72
          :::::||||||| ||||||||||||
Db 454 CTGCTAGAACTGCCACCTGAGAGAGTCATTGCCG 486

RESULT 11
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.:   6.57   Length: 4403765
Score: 122.50   Matches: 63
Percent Similarity: 36.53%   Conservative: 36
Best Local Similarity: 23.25%   Mismatches: 107
Query Match: 8.43%   Indels: 65
DB: 3   Gaps: 10

US-09-889-733B-2 (1-285) x US-09-103-840A-2 (1-4403765)

Qy   37 GluValAspValAlaArgAla------GluArgGlnHisGlnLeuTyrValGlyValLeu 54
          ::||| ||| ||||:::
...

Db 2591873 GACGTCCAAGTCGCGCAAGACCAGTGGGAGCACCTCCACCAGACCTATCTT------- 2591823

Qy   55 GlySerLysLeuGlyLeuGluValValGlyValGluLeuProAlaAspGluSerLeuProAspCys 74
          :::::|||| |||| :::
Db 2591822 ------CCGCTAGGCACCAGCGTGGATCTGATCGAGCGCCATTCCGGGTTACCGGACATG 2591769

Qy   75 ValPheValGluAspValAlaValCysGluGluThrAlaLeuIleThrArgProGly 94
          |||||| :::   ::::::
Db 2591768 GTGTACACCGCCACGGTGGGTTCATCGCGCAGACATCGCCGTGCGCCGGTTCCGG 2591709

Qy   95 AlaProSerArgArgLysGlyGluValAspMetMetLysGluAlaLeuGlyLysLeuGlnLeu 114
          :::   :::
Db 2591708 TTCCCGGACGAGCTGGTGAGTCCAGCTGGATGTCCTGGGTCCGGATAT 2591649

Qy 115 AsnIleValGluMetLysAspGluAsnAlaThrLeuAspGly---GlyAspValLeuPhe 133
          :|| ||| :::
Db 2591648 CGCCCGGTGACCACCGCCACGTCAAC--------GAGGACAGGGGGACCTGCTGATG 2591598

Qy 134 ThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGly---AlaGlu 152
          :::   ::::::
Db 2591597 GTTGGCGAAAGGGTGTTGGCGGCTACGGCTTTCGCACAGACCAGCGCGCACACGCCGAA 2591538

Qy 153 IleLeuAla--- 155
          :::||||||
Db 2591537 ATCGCCGCGTGCTTGGTCTGCCGGTGTGTGCCCGAGTTGGTGGTCGACCACGGTTCTAT 2591478

Qy 156 -----AspThr----------PheLysAspTyrArgAlaValSerThrValProValAla 169
              :::||| ||||||
Db 2591477 CACCTGGACACACCGGCCGCTGGCCGTGCTCGACGACCACCAGCGCCTACTACCGGCCGGCG 2591418

Qy 170 AspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGly 189
          :::   :::
Db 2591417 TTCAGTACGGCAGCGCAGGAACAGTTGCTGGGCGTGTTCCCGACGGATTGTGGTCGGC 2591358

Qy 190 SerSerGluSerAlaGlnHisLysAlaLeuGlyIleMetGlnGlnMetSerAspHisArgTyr 209
          :::::|||| :::
Db 2591357 AGTGCCGACGGCGTTCGTGTTCGGACTCAACGCCGGTCTCT----------------- 2591319

Qy 210 AspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsn 229
          :::|||| :::||||
Db 2591318 GACGGCGTCTGAACGTAGTGCTTCCGGTCGGGCCATGGGTTTTGCGGGCGCAGTTACGCGCA 2591259

Qy 230 LysGlyHisValLeuHisArgThrProGluThrProGluSerAlaLysValTyr 249
          :::   :::TTC 2591250
Db 2591258 GACGGTCTGAACGTAGTGCTTCCGGTCGGGCCATGGGTTTTGCGGGCGCAGTTACGCGCA ...

Qy 250 GluLysLeuLysAspHisMetLeuIleProValSerMetSerGluGluLuLysValAsp 269
          :::   ::::::
Db 2591249 GAGCCG------------------------GTCGGGTCGATCTGTCCGAGCTGCTCAAGGGCGGC 2591208

Qy 270 GlyLeuLeuThrCysCysSerValLeuIleAsn 280
          |||||::|||::::
Db 2591207 GGTTCCGTCAAGTGCTGCACCGTGCAGGAGATACAC 2591175

RESULT 12
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA

; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.:   93.1   Length: 4403765
Score: 113.50   Matches: 65
Percent Similarity: 34.83%   Conservative: 36
Best Local Similarity: 22.41%   Mismatches: 98
Query Match: 7.81%   Indels: 91
DB: Gaps: 11

US-09-889-733B-2 (1-285) x US-09-103-840A-2 (1-4403765)

Qy   59 GlyLeuGlnValValGluLeuProAlaAsp------------------------ 68
                |||::: |||||||||||||||||||||||||
Db 1117574 GGCATGAGACCTTCAACGAGCTCCCGTGGACACGCGGACGTGTCGTTGTGTTTGCGT 1117633

Qy   69 ---------------------GluSerLeuProAspCysValPheValGluAsp 79
                                  :::|||||
Db 1117634 ATGCACCATGGCGGAGACTTCGTCATTGAGCGTTGCGAACCTGGTGTTCACCCGCGAC 1117693

Qy   80 ValAlaValCysGluGluThrArgAlaLeuIleThrArgProSerArgArg 99
              ::::||| |||:::||
Db 1117694 TCGTCGATATGGATCGGGCCGGGTGGTCATTACGGGACCACGGGTG 1117753

Qy   100 LysGluValAspMet-----------------------------MetLys 106
              :::||| :::
Db 1117754 CGGGAAGCGTCGTGACCGACCTCATCTATGCTCATCACCCGGTTCACCCGTGTCGG 1117813

Qy   107 GluAlaLeuLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu 126
                                                              :::
Db 1117814 CGTGCCTATGAA--------------------------TCGGCACCGTCCGGTC 1117843

Qy   127 AspGlyGlyTyrAspValPheThrGlyArgGluPhe---ValGlyLeuSerLysArg 145
              :::|| |||::: :::|||
Db 1117844 GAGGGTGGCGACGTGTTGTTGCTCGCCCCGGGTGTGTCGGATGGGCGGAGCGGG 1117903

Qy   146 ThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp------------Tyr 161
              :::|||| |||||
Db 1117904 ACTACACCAGCAGCGCGGAAGCATTGGCGCGCAGCCTTTTGACGATGATCTTGCGCAT 1117963

Qy   162 AlaValSerThrValProValAlaAsp-----------GlyLeuHisLeuLysSerPheCys 178
              :::||| :::|||:::||
Db 1117964 ACCGTGCTCGCCGTGCGATCGTCAGCAGCGCGCAAATGCATCTCGACACGGTGTGC 1118023

Qy   179 SerMetAlaGly-- 182
              :::|||
Db 1118024 ACCATGGTCGACACGATACGGATGGTAGCCAACGTGTCGACACGCTCGAGGCG 1118083

Qy   183 ------------------------ProAsnLeuIleAlaIleGlySer------SerGlu 192
                                       |||||||
Db 1118084 TTCACGATCCAGCGCCACACCCGACCATCGGCGATGCGGCCCCGTTCGCCGAG 1118143

RESULT 13
US-09-103-840A-2
; Sequence 2, Application US/09103840A

; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION:
US-09-103-840A-1

Alignment Scores:
Pred. No.:   6.58   Length: 4411529
Score: 122.50   Matches: 63
Percent Similarity: 36.53%   Conservative: 36
Best Local Similarity: 23.25%   Mismatches: 107
Query Match: 8.43%   Indels: 65
DB: Gaps: 10

US-09-889-733B-2 (1-285) x US-09-103-840A-1 (1-4411529)

Qy   37 GluValAspValAlaArgAla------GluArgGlnHisLeuTyrValGlyValLeu 54
                ::::||||| |||:::
Db 2596063 GACGTCCAAGTCCGCAAGCACAGTGGGAGCACCTCCACCAGACCTATCTT------- 2596013

Qy   55 GlySerLysLeuGlyLeuGlnValValGluValGlyLeuProAlaAspGluSerLeuProAspCys 74
                :::|||||| |||
Db 2596012 ------CGGCTAGGCCACGGTGGATCTGAGCCCATTCCGGGTTACCGGACATG 2595959

Qy   75 ValPheValGluAspValAlaValCysGluGluThrArgAlaLeuIleThrArgProGly 94
              :::||| ::: |||:::::::
Db 2595958 GTGTACACCGCCAACGGTGGGTTCATCGCCGCACGACATCGCCCGGGTTCCGG 2595899

Qy   95 AlaProSerArgArgLysGluValAspMetMetLysGluAlaAlaGluLeuLysLeuGlnLeuLeu 114
              :::|| :::::: |||
Db 2595898 TTCCCGAACAGAGCTGGTGAGTCTAGAGCCTATGCCAGCTGATGTCCTCCGGTCGGATAT 2595839

Qy   115 AsnIleValGluMetLysAspGluAsnAlaThrLeuAspGly---GlyAspValLeuPhe 133
              :::||| :::|||:::|||
Db 2595838 CGCCCGGTGACCACCCGCACGTCAAC----------GAGGGACACAGGGCGACCTGCTGATG 2595788

Qy   134 ThrGlyArgGluPheValGlyLeuSerLysArgThrAsnGlnArgGly---AlaGlu 152
              :::||| :::||||
Db 2595787 GTTGGCGAAAGGGGTTGGCGGGCTACGGCTTTCGCACAGACAGCGGCACACGCCGAA 2595728

Qy   153 IleLeuAla-- 155
              :::|||
Db 2595727 ATCGCGCGCGGTGCTTGGTCTCGCCGGTGTCTCCCTCGAGTTGGTCGACCCACGGTTCTAT 2595668

Qy   156 -----AspThr-----------------PheLysAspTyrAlaValSerThrValProValAla 169
                    :::|| |||:::
Db 2595667 CACCTGGACACCGCGGCTGCCGTGCCGGTCCGTGCTACTACCCGCCGGCG 2595608

Qy   170 AspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGly 189
              :::|| |||
Db 2595607 TTCAGTACGGCAGCCGAGGAACAGTGTCGGCCGTGTTCCCGACGCCGATTGTGGTCGGC 2595548

Qy   190 SerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyr 209
              :::::: |||
Db 2595547 AGTGCCGACGGCGGTTCGTGTTCGGACTCAACGCCGTCT--------------- 2595509

Qy   210 AspLysLeuThrValProAspAspIleAlaLeuAlaAsnCysIleTyrLeuAsnIleProAsn 229
              :::||| :::
Db 2595508 GACGGTCTGAACGTAGTGCTTCCGTGCGCGCCAGTTACGGCCA 2595449

Qy   230 LysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyr 249
              :::||| |||
Db 2595448 GCCGGC---TTC 2595440

Qy   250 GluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAsp 269
              :::|| :::|||:::||
Db 2595439 GAGCCG------GTCGGTGTCGATCGTGTCCGAGCTGCTCAAGGGGCGGC 2595398

Qy   270 GlyLeuLeuThrCysCysSerValLeuIleAsn 280
              :::|||
Db 2595397 GGTTCGTCAAGTGCGCACGCTGGAGATACAC 2595365

```
Qy     193  SerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeu  212
                 :::||||    |||||::||    ||||||
Db  1118144  GCGGCTGCCAAGGCCATGGGAAATC----------------GACAAGCTG  1118176

Qy     213  ThrVal----------ProAspIleAlaAlaAsnCysIleTyrLeuAsnIlePro  228
                 |||           :::           |||
Db  1118177  CGGGTAATTCATACCGGAATGACACCCGTCGTCGGACGCGGACACGGC  1118236

Qy     229  AsnLysGlyHisValLeuHisArgThrProGluGluTyrProGluSerAlaLysVal  248
Db  1118237  AACAACACGTTGGCGTTGGCGCCCGGTGTCGTTCGCCTACGAGCGCAACGTACAGACC  1118296

Qy     249  TyrGluLysLysAspHis------MetLeuIleProValSerMetSerGluLeuGlu  266
                 :::||||::||       ::::::    ||||||
Db  1118297  AACGGCCCGCTGCAGGACGCGGGCATCGAAGTGCTTACCATCGCCCGGCTCCGAATTGGGT  1118356

Qy     267  LysValAspGlyLeuLeuThrCysCysSer  276
Db  1118357  ACCGGCCGCTGGCGGGCCCCGCTGCATGTCC  1118386
```

RESULT 14
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```
Alignment Scores:
Pred. No.:                          Length:          4411529
Score:                93.4          Matches:              65
                    113.50          Conservative:         36
Percent Similarity:  34.83%         Mismatches:           98
Best Local Similarity: 22.41%       Indels:               91
Query Match:          7.81%         Gaps:                 11
DB:                      3
```

US-09-889-733B-2 (1-285) x US-09-103-840A-1 (1-4411529)

```
Qy      59  GlyLeuGlnValValGlnGluLeuProAlaAsp-----------------  68
                 ||||:::          |||||::::|||
Db  1117552  GGCATGACCTTCAACGAGCTCCCGTCGACACGCGGACGACGTGTCGTTGGTCGTTGCGT  1117611

Qy      69  ------GluSerLeuProAspCysValPheValGluAsp  79
                       ||||:::          ||||::::|||
Db  1117612  ATGCACCATGGCGGAGACTTCGTCATTGAGCCGTTGCCGAACCTGGTGTTCACCCGCGAC  1117671

Qy      80  ValAlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArg  99
                 :::                         ||||
Db  1117672  TCGTCGGATATGGATCGGGCCGCGGGTGGTGATCCGCGGCTGGCATTACGGCACCGGGTG  1117731

Qy     100  LysGluValAspMet---------MetLys  106
                 :::::::
Db  1117732  CGGGAAGGCGTCGCTGACCGACCTCATCTATGCTCATCACCCCGGTTCACCGGTGTGCGG  1117791

Qy     107  GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu  126
                 |||:::                         ::::|||
Db  1117792  CGTGCCATGAA---------------------TCGCGGCACCGCTCCGGTC  1117821
```

```
Qy     127  AspGlyGlyAspValPheThrGlyArgGluPhePhe---ValGlyLeuSerLysArg  145
                 :::||||||||||           |||||:::||    :::
Db  1117822  GAGGGTGGCGACGTGTTGTTGCTCGCCCGGGTGGTCGCTGTCGGAGTGGGGGAGCGG  1117881

Qy     146  ThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp-------Tyr  161
                 ||||||           |||||:::          :::
Db  1117882  ACTACACCAGCAGGCGCGGAAGCATTGGCGCGACCTTTTGACGATGATCTTGCGCAT  1117941

Qy     162  AlaValSerThrValProValAlaAsp------GlyLeuHisLeuLysSerPheCys  178
Db  1117942  ACCGTGCTCGCCGTGCCATCGCTCAGCAGCCGCGCAAATGCATCTGACACGGTGTGC  1118001

Qy     179  SerMetAlaGly------  182
                 :::||
Db  1118002  ACGATGGTCGACACCGATACGATGGTGATGTGTACGCCAACGTTGTCGACACGCTCGAGGCG  1118061

Qy     183  ------ProAsnLeuIleAlaIleGlySer------SerGlu  192
                       |||:::          :::           :::||
Db  1118062  TTCACGATCCAGCGACACCCGACCATCGGCGCATGCGGCCCGTTCGCGGAG  1118121

Qy     193  SerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeu  212
                 :::||||    |::           |||||||
Db  1118122  GCGGCTGCCAAGGCCATGGGAAATC----------------GACAAGCTG  1118154

Qy     213  ThrVal----------ProAspIleAlaAlaAsnCysIleTyrLeuAsnIlePro  228
                 |||           :::           |||
Db  1118155  CGGGTAATTCATACCGGAATGACACCCGTCGTCGGACGCGGACACGGC  1118214

Qy     229  AsnLysGlyHisValLeuHisArgThrProGluGluTyrProGluSerAlaLysVal  248
                 |||||
Db  1118215  AACAACACGTTGGCGTTGGCGCCCGGTGTCGTTCGCCTACGAGCGCAACGTACAGACC  1118274

Qy     249  TyrGluLysLysAspHis------MetLeuIleProValSerMetSerGluLeuGlu  266
                 :::||||:::||       ::::::    ||||||
Db  1118275  AACGGCCCGCTGCAGGACGCGGGCATCGAAGTGCTTACCATCGCCCGGCTCCGAATTGGGT  1118334

Qy     267  LysValAspGlyLeuLeuThrCysCysSer  276
                 |||
Db  1118335  ACCGGCCGCTGGCGGGCCCCGCTGCATGTCC  1118364
```

RESULT 15
US-08-605-150A-7
; Sequence 7, Application US/08605150A
; Patent No. 6103520
; GENERAL INFORMATION:
; APPLICANT: Topfer, Reinhard
; APPLICANT: Hausmann, Ludger
; APPLICANT: Schell, Jozef
; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klein & Szekeres
; STREET: 4199 Campus Drive, Suite 700
; CITY: Irvine
; STATE: CA
; COUNTRY: USA
; ZIP: 92715
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,150A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT\EP94\02936
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P4329827.3
; FILING DATE: 03-SEP-1993

```
;.   ATTORNEY/AGENT INFORMATION:
;.     NAME: Szekeres, Gabor L.
;.     REGISTRATION NUMBER: 28,675
;.     REFERENCE/DOCKET NUMBER: 542-04-PA
;.   TELECOMMUNICATION INFORMATION:
;.     TELEPHONE: 714-854-5502
;.     TELEFAX: 714-854-4897
;. INFORMATION FOR SEQ ID NO: 7: -
;.   SEQUENCE CHARACTERISTICS:
;.     LENGTH: 1390 base pairs
;.     TYPE: nucleic acid
;.     STRANDEDNESS: double
;.     TOPOLOGY: linear
;.   MOLECULE TYPE: cDNA to mRNA
;.   HYPOTHETICAL: NO
;.   ANTI-SENSE: NO
;.   ORIGINAL SOURCE:
;.     ORGANISM: Cuphea lanceolata
;.   IMMEDIATE SOURCE:
;.     LIBRARY: ZAP cDNA library
;.     CLONE: C1GPDH30
;.   FEATURE:
;.     NAME/KEY: CDS
;.     LOCATION: 34..1149
US-08-605-150A-7

Alignment Scores:
Pred. No.:           0.00587        Length:        1390
Score:              102.00          Matches:         70
Percent Similarity:  37.50%         Conservative:    38
Best Local Similarity: 24.31%       Mismatches:     100
Query Match:          7.02%         Indels:          80
DB:                   3             Gaps:            16

US-09-889-733B-2 (1-285) x US-08-605-150A-7 (1-1390)

Qy     10 PheGlyArgAlaThrHisAlaValValArgAlaLeuProGluSerLeu---CysGlnHis  28
              |||||||      :::    ||||||      |||      ||||::|||
Db      4 TTCGGCACGAGTTCTTCTCAGCCTCGCAATGGCTCCCCTCGAGCTCAACTGCACCCAC  63

Qy     29 AlaLeuArgSerAlaLysGlyGluGluValAspValAlaArgAlaArgGlnHisGln  48
              ||||||||         :::
Db     64 CAGAACCCACATTCAAGCGGTTACGACGGACCCCAGATCGAGG------------  105

Qy     49 LeuTyrValGlyValLeuGlySer---LysLeuGlyLeuGlnValValGluLeuProAla  67
              ||||::|  |||||::||||||         |||    |||  ||||
Db    106 ------GTCACCGTTGTCGGTAGTGGAAACTGGGCGCAGTGTCGCTGCCAAGCTCATTGCT  159

Qy     68 AspGluSerLeuProAspCysValPheValGlyValAspValAlaVal---ValCysGluGlu  86
              |||      :::||||      :::|||      |||:::|||      ||||
Db    160 TCCAATACCCTCAAGCTTCCATCTTTTCATGATGAAGTGGGTATTTGAGGAG  219

Qy     87 ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLys  106
              ||||||         |:::::::
Db    220 ACT--------------CTACCGAGCGGCCAGAAGCTTACTGATGTCATCAAC  258

Qy    107 GluAlaLeuGlu----------------------LysLeuGlnLeuAsnIleValGlu  118
              :::                               |||:::::
Db    259 CAGACCAATGAAAATGTTAAGTATCTCCCGGAATTAAGCTCGGTAGGAATTTGTTGCA  318

Qy    119 MetLysAsp---GluAsnAlaThrLeuAspGlyGlyValAspPhe---ThrGlyArg  136
              |||      ::|||||      :::::||||
Db    319 GATCCAGACCTCGAAAACGCAGTTAAGGATGCAAATATGCTCGTGTTTGTGACACCGCAT  378

Qy    137 GluPhePheValGlyLeuSerLysArg--------ThrAsnGlnArgGlyValAlaGluIle  153
              ::|||||  |||:::   ||||||          ||||:::
Db    379 CAGTTCATGGAGGGCATCGCAAAAGACTCGTAGGAGAAAATACAGGAGAAGGAGCACAGGCT  438

Qy    154 LeuAlaAspThrPheLysPheAspTyrAlaValSerThrValProValAlaAspGlyLeuHis  173
              :::                               |||:::
Db    439 CTCTCC------------------------------CTTATAAAGGGCATCGAG  462

Qy    174 LeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerGlySer  193
```

```
Db    463 :::|||
              ::::|||   -----------ATGGAGGGGCCTTGCATGATCTCG--------  492
              |||  ||||||   :::|||

Qy    194 AlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisArgTyrAspLysLeuThr  213

Db    493 ----------------------------------------AGCCTA  498

Qy    214 ValProAspIleAlaAlaAsnCysIleTyrLeu----------AsnIleProAsnLys  230
              :::    |||   |||||      |||||||:::
Db    499 ATCTCTGATCTTCTCGGGATCAACTGCGTGTCCTAATGGGGCAAACATCGCTAATGAG  558

Qy    231 GlyHisValLeuLeuHisArgThrProGluGlu--------TyrProGluSerAlaLys  247
              :::  :::  :::      |||:::
Db    559 ------ATTGCTGTTGAGAAATTCAGTGAAGCGACAGTCGGGTTCAGAGAAAAATACAGAT  612

Qy    248 ValTyrGluLysLeuLysAspHisMetLeuIleProValSerMet---SerGluLeuGlu  266
              :::
Db    613 ATTGCGGAGAAATGGGTTCAGCTCTTTAGCACTCCGTACTTCATGGTCTTCAGCTGTTGAA  672

Qy    267 LysValAspGlyLeuLeuThrCys  274
              |||:::::|||||
Db    673 GATGTTGAGGAGTAGAACTTTGT  696

Search completed: June  8, 2004, 03:30:29
Job time : 8710 secs
```

```
                    GenCore version 5.1.6
           Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:   June 7, 2004, 23:21:06 ; Search time 389 Seconds
                                    (without alignments)
                      3342.336 Million cell updates/sec

Title:          US-09-889-733B-2
Perfect score:  1454
Sequence:       1 MAGLGHPSAFGRATHAVVRA........EKVDGLLTCCSVLINKKVDS 285

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       295936 seqs, 2280998010 residues

Total number of hits satisfying chosen parameters:   5991872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match   0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09889733/runat_07062004_083529_8024/app_query.fasta_1.455
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09889733 @CGN_1_723 @runat_07062004_083529_8024
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA.*
          1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
          2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
          3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
          4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
          5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
          6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
          7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
          8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
          9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
         10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
         11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
         12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
         13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
         14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
         15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
         16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
         17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
         18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
         19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Flag	Result No.	Score	Query Match	Match Length	DB	ID	Description
	1	1439	99.0	1633	17	US-10-450-826-9	Sequence 9, Appli
	2	1439	99.0	4208	17	US-10-468-125-22	Sequence 22, Appl
	3	1379	94.8	3008	15	US-10-205-219-83	Sequence 83, Appl
	4	1200	82.5	3778	15	US-10-252-157-328	Sequence 328, App
	5	1064	73.2	3651	9	US-09-925-300-512	Sequence 512, App
	6	645.5	44.4	1397	17	US-10-641-643-37	Sequence 37, Appl
	7	551.5	37.9	770	9	US-09-910-943-9	Sequence 9, Appli
	8	386	26.5	472	10	US-09-918-995-9117	Sequence 9117, Ap
	9	382	26.3	452	10	US-09-918-995-14871	Sequence 14871, A
	10	342	23.5	403	10	US-09-918-995-19254	Sequence 19254, A
	11	282.5	19.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
	12	281	19.3	774	15	US-10-156-761-1679	Sequence 1679, Ap
	13	266	18.3	332	9	US-09-783-590-5309	Sequence 5309, Ap
	14	265.5	18.3	83946	17	US-10-450-826-8	Sequence 8, Appli
c	15	233	16.0	239	9	US-09-736-457-404	Sequence 404, App
c	16	233	16.0	239	9	US-09-902-941-404	Sequence 404, App
c	17	233	16.0	239	9	US-09-849-826-404	Sequence 404, App
c	18	233	16.0	239	10	US-09-476-300-404	Sequence 404, App
c	19	233	16.0	239	13	US-1-283-017-404	Sequence 404, App
c	20	233	16.0	239	15	US-10-017-754-404	Sequence 404, App
c	21	233	16.0	239	15	US-10-113-872-404	Sequence 404, App
c	22	204	14.0	3673778	15	US-10-312-841-1	Sequence 1, Appli
c	23	198	13.6	3189	10	US-09-814-353-20553	Sequence 20553, A
c	24	165	11.3	3673778	15	US-10-312-841-2	Sequence 2, Appli
	25	153.5	10.6	475	10	US-10-918-995-16593	Sequence 16593, A
	26	117	8.0	1377	13	US-10-282-122A-17396	Sequence 17396, A
	27	107.5	7.4	1218	15	US-10-238-075-714	Sequence 714, App
	28	107.5	7.4	1349	16	US-10-369-493-38269	Sequence 38269, A
	29	107.5	7.4	1785	16	US-10-369-493-38603	Sequence 38603, A
	30	107.5	7.4	1821	16	US-10-369-493-35077	Sequence 35077, A
	31	107.5	7.4	3943	16	US-10-398-221-3727	Sequence 3727, Ap
c	32	107.5	7.4	6164	15	US-10-238-075-705	Sequence 705, App
c	33	107.5	7.4	6425	15	US-10-085-959-217	Sequence 217, App
	34	105.5	7.3	1059	16	US-10-369-493-37878	Sequence 37878, A
	35	103.5	7.1	1227	15	US-10-156-761-2307	Sequence 2307, Ap
c	36	103.5	7.1	9025608	15	US-10-156-761-1	Sequence 1, Appli
	37	102.5	7.0	1827	10	US-09-738-269-40	Sequence 40, Appl
	38	102.5	7.0	1827	14	US-10-023-437-40	Sequence 40, Appl
c	39	102.5	7.0	684707	16	US-10-398-221-9	Sequence 9, Appli
	40	102.5	7.0	3011208	16	US-10-398-221-2058	Sequence 2058, Ap
c	41	102	7.0	163	15	US-10-029-386-26699	Sequence 26699, A
c	42	102	7.0	528	15	US-10-029-386-12999	Sequence 12999, A
	43	102	7.0	1230	16	US-10-369-493-40383	Sequence 40383, A
c	44	101.5	7.0	135638	15	US-10-314-657-1	Sequence 1, Appli
	45	101	6.9	2832	13	US-10-282-122A-31868	Sequence 31868, A

ALIGNMENTS

```
RESULT 1
US-10-450-826-9
; Sequence 9, Application US/1045026
; Publication No. US20040101818A1
; GENERAL INFORMATION:
;  APPLICANT: Ji, Darren
;  APPLICANT: Axelrod, Douglas W.
;  APPLICANT: Cook, Jonathon S.
;  APPLICANT: Jaiswal, Neelam
;  APPLICANT: Eistein, Richard
;  APPLICANT: Houghton, Adam
;  APPLICANT: Mertz, Lawrence
;  TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiatio
;  FILE REFERENCE: 044921-5039-WO
;  CURRENT APPLICATION NUMBER: US/10/450,826
;  CURRENT FILING DATE: 2003-06-18
;  PRIOR APPLICATION NUMBER: US 60/255,882
;  PRIOR FILING DATE: 2000-12-18
;  PRIOR APPLICATION NUMBER: US 60/285,691
;  PRIOR FILING DATE: 2001-04-24
;  NUMBER OF SEQ ID NOS: 149
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
```

; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AB001915
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(1633)
; OTHER INFORMATION: n = a or c or g or t
US-10-450-826-9

Alignment Scores:
Pred. No.:   2.21e-177
Score: 1439.00
Percent Similarity: 99.65%
Best Local Similarity: 99.30%
Query Match: 98.97%
DB:

Length: 1633
Matches: 283
Conservative: 1
Mismatches: 1
Indels: 0
Gaps: 0

US-09-889-733B-2 (1-285) x US-10-450-826-9 (1-1633)

Qy   1   MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla   20
Db   323 ATGGCCGGGCTCGGCCACCCGAGCGCCTTCGGCCGCGGCCACCCAGCCGTGGTGCGGGCG 382

Qy   21   LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValValAspVal   40
Db   383 CTACCCGAGTCGCTGTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGTGGACGTC 442

Qy   41   AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuSerLeuGlyLeu   60
Db   443 GCCCGCGCGGAACGGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCGTG 502

Qy   61   GlnValValGluLeuProAlaAspCysValPheGlyValGlnAspVal   80
Db   503 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGGTGAGGACGTG 562

Qy   81   AlaValValCysGluGluThrAlaIleuIleThrArgProGlyAlaProSerArgArgLys 100
Db   563 GCCGTGGTGTGCGAGGAGACGACGGGCCCTCATCACCCGACCCGGGGCGCCGAGGGAAG 622

Qy 101   GluValAspMetLysGluLysLeuGlnLeuLeuGluAsnIleValGlyMetLys 120
Db   623 GAGGTTGACATGATGAAAGAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 682

Qy 121   AspGluAsnAlaThrLeuAsnGlyGlyAspValGlyAspPheArgGluPhePheVal 140
Db   683 GATGAAAATGCAACTTAGATGGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTG 742

Qy 141   GlyLeuSerLysArgThrArgGlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysAsp 160
Db   743 GGCCTTTCCAAAAGGACAAATCAACGAGGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 802

Qy 161   TyrAlaValSerThrValProValAlaAlaspGlyLeuHisLeuLysSerSerMet 180
Db   803 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 862

Qy 181   AlaGlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIle 200
Db   863 GCTGGGCCTAACCTGATCGCAATTGGGTCTGAGTCAGTGCTCCACAGGCCCTTAAGATC 922

Qy 201   MetGlnGlnMetSerAspHisArgTyrAspLysTyrLeuThrValProAspAspIleAlaAla 220
Db   923 ATGCAACAGATGAGTGACCACGTGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 982

Qy 221   AsnCysIleTyrLeuAsnIleProAsnLysGlyHisHisLeuLeuHisArgThrProGlu 240
Db   983 AACTGTATATATCTAAATATCCCACAAAGGACGCACGTTCTTGCTCCACCGAACCCGGAA 1042

Qy 241   GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
Db 1043 GAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCACATGCTGATCCCCGTG 1102

Qy 261   SerMetSerGluLeuGluLysValAlaAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
Db 1103 AGCATGTCTGAACTGGAAAAGGTGGAAGGGCTGCTCACCTGCTGCTCAGTTTTAATTAAC 1162

Qy 281   LysLysValAspSer 285
Db 1163 AAGAAGTAGACTCC 1177

RESULT 2
US-10-468-125-22
; Sequence 22, Application US/10468125
; Publication No. US20040082061A1
; GENERAL INFORMATION:
;   APPLICANT: ASTROMOFF, Anna
;   APPLICANT: AU-YOUNG, Janice
;   APPLICANT: BAUGHN, Mariah R.
;   APPLICANT: DING, Li
;   APPLICANT: DUGGAN, Brendan M.
;   APPLICANT: FORSYTHE, Ian J.
;   APPLICANT: GIETZEN, Kimberly J.
;   APPLICANT: GRIFFIN, Jennifer A.
;   APPLICANT: LEE, Ernestine A.
;   APPLICANT: LU, Yan
;   APPLICANT: RICHARDSON, Thomas W.
;   APPLICANT: RING, Huijun Z.
;   APPLICANT: SANJANWALA, Madhusudan
;   APPLICANT: SWARNAKAR, Anita
;   APPLICANT: CHAWLA, Narinder K.
;   APPLICANT: WARREN, Bridget A.
;   APPLICANT: XU, Yuming
;   APPLICANT: YUE, Henry
;   APPLICANT: ZEBARJADIAN, Yeganeh
;   TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
;   FILE REFERENCE: PI-0363 USN
;   CURRENT APPLICATION NUMBER: US/10/468,125
;   CURRENT FILING DATE: 2003-08-15
;   PRIOR APPLICATION NUMBER: PCT/US02/04918
;   PRIOR FILING DATE: 2002-02-14
;   PRIOR APPLICATION NUMBER: US 60/269,643
;   PRIOR FILING DATE: 2001-02-16
;   PRIOR APPLICATION NUMBER: US 60/271,332
;   PRIOR FILING DATE: 2001-02-23
;   PRIOR APPLICATION NUMBER: US 60/276,767
;   PRIOR FILING DATE: 2001-03-16
;   PRIOR APPLICATION NUMBER: US 60/282,077
;   PRIOR FILING DATE: 2001-04-06
;   PRIOR APPLICATION NUMBER: US 60/285,447
;   PRIOR FILING DATE: 2001-04-19
;   PRIOR APPLICATION NUMBER: US 60/287,060
;   PRIOR FILING DATE: 2001-04-27
;   PRIOR APPLICATION NUMBER: US 60/288,543
;   PRIOR FILING DATE: 2001-05-03
;   NUMBER OF SEQ ID NOS: 24
;   SOFTWARE: PERL Program
; SEQ ID NO 22
;   LENGTH: 4208
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   OTHER INFORMATION: Incyte ID No: 7671089CB1
US-10-468-125-22

Alignment Scores:
Pred. No.:   9.59e-177
Score: 1439.00
Percent Similarity: 99.65%
Best Local Similarity: 99.30%
Query Match: 98.97%
DB:

Length: 4208
Matches: 283
Conservative: 1
Mismatches: 1
Indels: 0
Gaps: 0

US-09-889-733B-2 (1-285) x US-10-468-125-22 (1-4208)

; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
;   LENGTH: 3008
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
; FEATURE:
;   OTHER INFORMATION: N-G, N-G dimethylarginine dimethylaminohydrolase
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 2411, 2412
;   OTHER INFORMATION: n is a or g or c or t
US-10-205-219-83

Alignment Scores:
Pred. No.:   3008
Score: 3.98e-169   Length: 3008
Percent Similarity: 1379.00   Matches: 268
Best Local Similarity: 96.84%   Conservative: 8
Query Match: 94.04%   Mismatches: 9
                       94.84%   Indels: 0
DB: 15   Gaps: 0

US-09-889-733B-2 (1-285) x US-10-205-219-83 (1-3008)

Qy   1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
           ||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db   432 ATGGCCGGGCTCGGCCACCCCTCAGCCACCCCTCGGTCTTCGGCCGGGCCACCCACGCCGTGGTGCGGGCT 491

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValLeuAspVal 40
           |||||||||||||||||||||||||||:::||||||||||||||||||||||
Db   492 CCGCCCGAGTCCCTGCCGCCACGCGCTCCCAGGGCAAGAGGTGGATTTC 551

Qy   41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuSerLysGlyLeuGlyLeu 60
           |||||||||||||||||||||||||||:::||||||||||||||||||||||
Db   552 GCTCGCCGCTGAGCGCCGCCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCTGGGGCTG 611

Qy   61 GlnValValGlyLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
           |||||||||||||||||||||||||||:::||||||||||||||||||||||
Db   612 CAGGTGGTGCAGCTGCCGGCCGACGAGAGCCTGCCTGACGTGGTGTTCGTGGAGGACGTG 671

Qy   81 AlaValValCysGluGluThrArgProGlyAlaProSerArgArgLys 100
           |||||||||||||||||||||||||||:::||||||||||||||||||||||
Db   672 GCCGTGGTGTGCGAGGAGACCCGCCCTGGAGGCGGGGCCCTAGCCGCCAGGAAG 731

Qy   101 GluValAspMetMetLysGluLeuGlnLeuAsnIleValGluMetLys 120
           |||||||||||||||||||||||||||:::||||||||||||||||||||||
Db   732 GAGGTTGACATGATGAAGGAGCTTTGGAAAACTTCAGTCAACATAGTAGAGATGAAA 791

Qy   121 AspGluAsnAlaThrLeuAspGlyGlyAspArgGlyAlaGluIleLeuAlaAspThrPheLysAspVal 140
           |||||||||||||||||||||||||||:::||||||||||||||||||||||
Db   792 GATGAAAATGCAACCTTAGATGGTGGGACGTCCTATTCACAGGCAGGAGTTTTTGTG 851

Qy   141 GlyLeuSerLysArgThrArgAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
           |||||||||||||||||||||||||||:::||||||||||||||||||||||
Db   852 GGCCTTCCAAAAGGACAAATCACCGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGAC 911

Qy   161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
           |||||||||||||||||||||||||||:::||||||||||||||||||||||
Db   912 TACGCAGTTTCCACAGTCCCCGTGGCCGATTCTTTGCATTTAAAGAGTTTCTGCAGCATG 971

Qy   181 AlaGlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIle 200
           |||||||||||||||||||||||||||:::||||||||||||||||||||||
Db   972 GCTGGGCCCCAACCTGATCGCAATAGGTCCAGAAAGGCCCTCAAGATC 1031

Qy   201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
           |||||||||||||||||||||||||||:::||||||||||||||||||||||
Db 1032 ATGCAACAGATGAGTGACCACCGTTATGACAAGCTCACTGTACCGGACGACATGGCCGCC 1091

Qy   221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
           |||||||||||||||||||||||||||:::||||||||||||||||||||||

Qy   20 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db   368 ATGGCCGGGCTCGGCCACCCCTCGGCCGCCTTCGGCCGGGCCACCCACGCCGTGGTGCGGGCG 427

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValLeuAspVal 40
Db   428 CTACCCGAGTCGCTGCCGCCACGCGCTCCGCCAGCAGCGGCCCAGGGCCAGGAGGTGGACGTC 487

Qy   41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuSerLysGlyLeuGlyLeu 60
Db   488 GCCCGCGCGGAACGGCAGCAGCTCTACGTGGGCGTGCTGGGCAGCAAGGGGCTG 547

Qy   61 GlnValValGlyLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db   548 CAGGTGGTGGGCCTGCCCGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 607

Qy   81 AlaValValCysGluGluThrArgProGlyAlaProSerArgArgLys 100
Db   608 GCCGTGGTGTGCGAGGAGACCCGCCCCCATCACCCGACCCGGGGGCCGAGCCGGAGGAAG 667

Qy   101 GluValAspMetMetLysGluLeuGlnLeuAsnIleValGluMetLys 120
Db   668 GAGGTTGACATGATGAAGGAGCTTCAATTAGAAAAAACTTCAGTCAATAGTAGAGATGAAA 727

Qy   121 AspGluAsnAlaThrLeuAspGlyGlyAspArgGlyAlaGluIleLeuAsp 140
Db   728 GATGAAAATGCAACCTTAGATGGCGGAGATGTTTATTCACAGGCAGAGAATTTTTGTG 787

Qy   141 GlyLeuSerLysArgThrArgAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
Db   788 GGCCTTTCCAAAAGGACAAATCAACAGGAGTGCTGAAATCTTGGCTGATACTTTTAAGGAC 847

Qy   161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db   848 TATGCAGTCTCCACAGTGCCAGTGGCCGATGGGTTGCATTTGAAAGAGTTTCTGCAGCATG 907

Qy   181 AlaGlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIle 200
Db   908 GCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 967

Qy   201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db   968 ATGCAACAGATGAGTGACCACCGTTACGACAAACTCACTGTGCCTGATGACATAGCAGCA 1027

Qy   221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 1028 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCCGGAA 1087

Qy   241 GluTyrProGluSerAlaLysValThrTyrGluLysLeuAspHisMetLeuIleProVal 260
Db 1088 GAGTATCCAGAAAGTGCAAAGGTTATGAGAAAACTGAAGGACCATATGCTGATCCCGTG 1147

Qy   261 SerMetSerGluLeuGluLysLeuValAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
Db 1148 AGCATGTCTGAACTGGAAAAGCTGGAAAAAGGTGGATGGGCTGCTTCACCTGCTTGCTTCAGTTTAATTAAC 1207

Qy   281 LysLysValAspSer 285
Db 1208 AAGAAAGTAGACTCC 1222

RESULT 3
US-10-205-219-83
; Sequence 83, Application US/10205219
; Publication No. US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brookbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219

```
Db    1092  AACTGTATATATTAAATATCCCAGCAAAGGGCATGTCTTGCTGCACCGAACCCCAGAA  1151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302  CCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCC  361

Qy     241  GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal  260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     167  oValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAl  187

Db    1152  GAGTACCCAGAAAGCGCAAAGGTTTATGAGAGCTCAAGAGACCATCTACTGATCCCTGTG  1211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362  AGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGC  421

Qy     261  SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn  280
             ||||||||||||||:::||||||||||||||||||||||||||||   ||||||||
Qy     187  aIleGlySerGluSerAlaGlnLysAlaLeuLysValIleMetGlnMetSerAspHi  207

Db    1212  AGCAATTCTGAGATGATGGAAAAGGTGGACGGGCTTGCTCACCTGCTGCTCCGTTTTTAC  1271
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422  AATTGGGTCTAGTTGAATCTGCACAGAGAGCCCTTAAGATCATGCAACAGATGAGTGACCA  481

Qy     281  LysLysValAspSer  285                              Qy     207  sArgTyrAspLysLeuThrValProAspAspAspIleAlaAlaAsnCysIleTyrLeuAsnIl  227
             |||||||||||||                                       ||||||||||||||||||||||||||||||||||||
Db    1272  AAGAAGACAGACTCT  1286                             Db     482  CCGCTACGACAACAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCTAAAATAT  541

RESULT 4                                            Qy     227  eProAsnLysGlyHisValLeuHisLeuHisArgThrProGluTyrProGluSerAlaIy  247
US-10-252-157-328                                          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::||
; Sequence 328, Application US/10252157           Db     542  CCCCAACAAAGGGCACGTCTTGCTGCACCGACCCCGAAGAGTAGTCCAGAAAGTGCAA  601
; Publication No. US20030190640A1
GENERAL INFORMATION:                               Qy     247  sValTyrGluLysLysAspHisMetLeuIleProValSerMetSerGluLeuGluLy  267
  APPLICANT: Paris, Mary                                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  APPLICANT: Pearson, Cecelia I.                   Db     602  GGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCGTGAGCATGTCTGAACTGGAAA  661
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US                     Qy     267  sValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer  285
; CURRENT APPLICATION NUMBER: US/10/252,157               ||||||||||||||||||||||||||||||||||||||||||||||||||
; CURRENT FILING DATE:  2002-10-01                 Db     662  GGTGGATGGGCTGCTCACCTGCTCTCCAGTTTTAATTAACAAGAAAGTAGAGACTCC  716
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31                    RESULT 5
; NUMBER OF SEQ ID NOS: 501                        US-09-925-300-512
; SOFTWARE: PERL Program                           ; Sequence 512, Application US/09925300
SEQ ID NO 328                                      ; Patent No. US20020151681A1
; LENGTH: 3778                                     GENERAL INFORMATION:
; TYPE: DNA                                          APPLICANT: Craig Rosen,
; ORGANISM: Homo sapiens                             APPLICANT: Steve Ruben
; FEATURE:                                         ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; NAME/KEY: misc_feature                           ; FILE REFERNCE: PA101
; OTHER INFORMATION: Incyte ID No. US20030190640A1 198087.1  ; CURRENT APPLICATION NUMBER: US/09/925,300
; FEATURE:                                         ; CURRENT FILING DATE:  2001-08-10
; NAME/KEY: unsure                                 ; PRIOR APPLICATION NUMBER: PCT/US00/05988
; LOCATION: 1083-1106                              ; PRIOR FILING DATE: 2000-03-08
; OTHER INFORMATION: a, t, c, g, or other         ; PRIOR APPLICATION NUMBER: 60/124,270
US-10-252-157-328                                  ; PRIOR FILING DATE: 1999-03-12
                                                   ; NUMBER OF SEQ ID NOS: 1890
Alignment Scores:                                  ; SOFTWARE: PatentIn Ver. 2.0
Pred. No.:                         Length:    3778 SEQ ID NO 512
Score:           1.43e-145         Matches:    238 ; LENGTH: 3651
Percent Similarity: 99.58%         Conservative:  0 ; TYPE: DNA
Best Local Similarity: 99.58%      Mismatches:    0 ; ORGANISM: Homo sapiens
Query Match:     82.53%            Indels:        1 ; FEATURE:
DB:                                Gaps:          0 ; NAME/KEY: misc_feature
                                                   ; LOCATION: (1283)
US-09-889-733B-2 (1-285) x US-10-252-157-328 (1-3778)  ; OTHER INFORMATION: n equals a,t,g, or c
                                                   ; NAME/KEY: misc_feature
Qy      48  GlnLeuTyrValGlyLeuValLeuGlySerLysLeuGlyLeuGlnValValGluProAla  67  ; LOCATION: (3641)
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||  ; OTHER INFORMATION: n equals a,t,g, or c
Db       2  CAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCGGCC  61  ; NAME/KEY: misc_feature
                                                   ; LOCATION: (3650)
Qy      68  AspGluSerLeuProAspCysProValPheValGluValAlaValAlaValCysGluThr  87  ; OTHER INFORMATION: n equals a,t,g, or c
             |||||||||||||||||||||||||||||||||||||||||||||||||||||  US-09-925-300-512
Db      62  GACGAGAGCCTTCCGGACTGCCCTGTCTTCGTGGAGGTCGCTGTGGCGGTGTGCGAGACG  121
                                                   Alignment Scores:
Qy      88  AlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLysGlu  107  Pred. No.:                        Length:    3651
             ||||||||||||||||||||||||||||||||||||||||||||  Score:         8.2e-128         Matches:    212
Db     122  GCCCTCATCACCCGACCCGGGCGCCCGAGCCCGGAGGAGGAAGGTTGACATGATGAAAGAA  181  Percent Similarity: 99.53%      Conservative:  0
                                                   Best Local Similarity: 99.53% Mismatches:    1
Qy     108  AlaLeuGluLysLeuGlnLeuAsnIleValGlnMetLysAspGluAsnAlaThrLeuAsp  127  Query Match:  73.18%            Indels:        1
             |||||||||||||||||||||||||||||||||||||||||||||||  DB:           9                 Gaps:          0
Db     182  GCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGAT  241
                                                   US-09-889-733B-2 (1-285)  x US-09-925-300-512 (1-3651)
Qy     128  GlyGlyAspValPheThrGlyPheArgGluPheValGlyLeuSerLysArgThrAsn  147
             ||||||||||||||||||||||||||||||||||||||||||||||||||  Qy      73  AspCysValPheValGluAspValAlaValValCysGluThrAlaLeuIleThrArg  92
Db     242  GGCGGAGATGTTTTTATTCACAGGCAGAGAATTTTGTGGGCCTTTCCAAAAGGACAAAT  301         |||||||||||||||||||||||||||||||||||||||||||||||
Qy     148  -GlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysAspThrPheLysSerThrValPr  167  Db       4  GACTCGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGTGAGACGGCGCCCATCACCCGA  63
```

REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (650) 855-0555
  TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 37:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1397 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  IMMEDIATE SOURCE:
    LIBRARY: HMC1NOT01
    CLONE: 003437
  SEQUENCE DESCRIPTION: SEQ ID NO: 37 :

US-10-641-643-37

Alignment Scores:
Pred. No.:   Length: 1397
Score: 9.58e-74   Matches: 141
Percent Similarity: 645.50   Conservative: 46
Best Local Similarity: 68.75%   Mismatches: 78
Query Match: 51.84%   Indels: 7
DB: 44.39%   Gaps: 3
                           17

US-09-889-733B-2 (1-285) x US-10-641-643-37 (1-1397)

Qy   3 GlyLeuGlyHisPro---SerAlaPheGlyAlaThrHisAlaValValArgAlaLeu 21
          ||| :||| :||::|:|||:|::|:||||||||:|| ||:||
Db   332 GGGATGGGACGCCGGGGAGGGGGCTGGGCCGCTGCTCCCATGCCCGATCCGGGGAGTC 391

Qy   22 ProGluSerLeuCysGlnHisAlaAlaLeuArgSerAlaLeuGlyGluValAspValAla 41
          || || ::|:|||| || ::||||:||:
Db   392 CCAGAGAGCCTTGGCGTTCGGGGGAAGTGCGGGGGGCTGGGCCTTCCCGCTCTTGGATCTCGGCC 451

Qy   42 ArgAlaArgGlnHisGlnLeuTyrValGlyValSerLeuGlyLeuGln 61
          |||| :|||||||| |: ||||| :|||||| |:|:|
Db   452 AAAGCTCAAAGGAGCACGGGGGTGCTGGGAGTAAACTGAGCAACGACTGGGGCTACAG 511

Qy   62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
          :|| :|| ||||
Db   512 CTGCTAGAACTGCCACCTGAGGAGTCATTGCGCGTGGGACCGCTGTTGGGCACACGGCC 571

Qy   82 ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
          ||:::| ||||:::|::||| || :||::|||
Db   572 GTGATCCAAGGGACACGGCCCTAATCACGGCGCCCGGGCCCCGGTCCGTCGTCAGGCCCAGAG 631

Qy   102 ValAspMetMetLysGluGluLysGluLeuGlnLeuAsnIleValGluMetLysAsp 121
          ||:| :::|||||:|:|:||:|:|:| |:|||||||
Db   632 GTCCGATGGAGTCCGCAAAGCCTCGAAGACCTGGGGCTCCGAATTGTGGAAATAGGAGAG 691

Qy   122 GluAsnAlaThrLeuAspGlyValGlyAspValAspValLeuPheThrGlyArgGluPheValGly 141
          |||::|:::|||| :|:||||:|:||||||:|:|:||:|
Db   692 GAGAACGCCGACCTGGATGGCACTGACGTTCTTCACCGGCCGGGAGAGTTTTTCGTAGGC 751

Qy   142 LeuSerLysArgThrAsnGlnArgGlyAla-GluLeuLeuAlaAspThrProLysAspTy 161
          |||||:|:::|||||||||||:||| |||||||
Db   752 CTCTCCAAATGACCAATCACCGAGGGCCTGCGAGATCGTGGCGGACACGTTCCGGGACTT 811

Qy   161 rAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAl 181
          :||||||||:|::||::|:|:||| :||:||||||
Db   812 CGCCGTCCCACTGTGCCAGTCTCGGGTCCCTCCCACCTGCGGCGTCTCGTGCGGCATGGG 871

Qy   181 aGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMe 201
          |||| :| |||||:||||||:
Db   872 GGGACCTTCGCACTGTTGTGCAGGCAGCAGGCGTGCCCAAAAGGCTGTCCGGGCAAT 931

Qy   201 tGlnGlnMetSerAspHisArgTyrAspLysLeuThr-ValProAspAspIleAlaAlaA 221
          ||||:|||:::||||||:|||||:|| |||:|:||:
Db   932 GGCAGTGCTGACAGATCACCCATATGCCTCCCAGATGACGCAGCGCATGCTG 991

Qy   221 snIleProAsnLeuTyrLeu-----AsnIleProAsnGlyHisValLeuHisLeuHisArgThrP 239
          ::|||:||||

Db   992 ACTGTCTCTTTCTTCGTCCTGGGTTGCCTGGTGTGCCCCTTTCCTCCTGCACCGTGAG 1051

Db   591 GCTGAAATGGTGCAAAGACCTACCAGGATTATGCTGTTTCAACTGTTCCTGTGTCTGGG 650

Qy   239 roGluGlu------TyrProGluSerAlaLysAlaValTyrGluLysLeuLysAspHisMetL 257
              ::: ||| :::   :||||||||||||
Db   1052 GTGGGGATCTGTGCNCCCAAACAGCCAGAGGAGCACTGAGGAAGCTCTCGATGATGTCACCC 1111

Qy   171 GlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySer 190
              ::::||| ||| ::: :::||
Db   651 GACATGCACTTTAAAAGCTTTGCAGTATGCAGGCACTGACACCCTGCTCATAGGAGAGC 710

Qy   257 euIleProValSerMetSerGluLeuGlnLys 267
              ||:|||||||||
Db   1112 TGGTACCTGTGTCCTGCTCAGAACTGGAGAGAG 1143

Qy   191 SerGluSerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHis 207
              |:::|| ||||||
Db   711 AGTGATACAGCGAGAAGGCACTTGAGACTATGGACAACTGACCGATCAT 761

RESULT 7
US-09-910-943-9
; Sequence 9, Application US/09910943
; Patent No. US2002081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Brivanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/1G14US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 9
; LENGTH: 770
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(770)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-9

RESULT 8
US-09-918-995-9117
; Sequence 9117, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9117
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-9117

Alignment Scores: Alignment Scores:
Pred. No.:   Pred. No.:
Score: 7.41e-62   Length: 770   Score: 1.51e-40   Length: 472
Percent Similarity: 551.50   Matches: 110   Percent Similarity: 386.00   Matches: 83
Best Local Similarity: 74.11%   Conservative: 36   Best Local Similarity: 72.00%   Conservative: 25
Query Match: 55.84%   Mismatches: 46   Query Match: 55.33%   Mismatches: 30
DB: 37.93%   Indels: 5   26.55%   Indels: 12
                                  Gaps: 1   Gaps: 1
                                                     DB: 10

US-09-889-733B-2 (1-285) x US-09-910-943-9 (1-770)   US-09-918-995-9117

                                                     US-09-889-733B-2 (1-285) x US-09-918-995-9117 (1-472)

Qy   11   GlyArgAlaThrHisAlaValValArgAlaLeuProGluSerLeuCysGlnHisAlaLeu 30
             |||||| ||| :::||
Db   186 GGGCGCTACACTGCACGCTGTGTGAGGGTGTCCCCTCGTCCTTAGAAAAAGAGACC--- 242

Qy   30   LeuArgSerAlaLysGlyGluGlu------------------------------ 37
             ::::||||
Db   21   CTGAGACTCTCGGAAGCGCCGGAATTCGGCGGGGGAAGGTGCGGGGGCTGGCCTTCCCGCT 80

Qy   31   ArgSerAlaLysGlyGluGluValAspValAlaArgValAlaArgAlaAlaGluArgGlnHisGlnLeuTyr 50
             :::|||| :::||| :::
Db   243 ------------AATGGTCAGGTAGACCTGGCACCGGGCTCAACGTGAGAATGGAGTCTAC 290

Qy   38   ValAspValAlaArgAlaAlaGlnHisGlnLeuTyrValGlyValLeuGlyValValGlySerLys 57
             |||||| ::: |||
Db   81   CTGGATCTGGCCAAAGCTCAAAGGGAGCACGGGGGGCTGGGAGTAAACTGAGCAACGA 140

Qy   51   ValGlyValLeuGlySerLysLysGluGlyLysLeuGlnValValGluGluLeuProAlaAspGluSer 70
             ||:||| ||:::
Db   291 TGTGGCATCCTGAGACAGAAACTTGGGCTGCGAGGTGGTGCAGGTGGTGCCCCCCAATGAGGAA 350

Qy   58   LeuGlyLeuGlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheVal 77
             ||||| ::: |||||
Db   141 CTGGGGCTACAGCTGCTAGAACTGCCACCTGAGGAGGCATTGCCGTGGGACCCGTGCTT 200

Qy   71   LeuProAspCysValPheValGlnHisValAspSerValAlaValAlaValCysGluGluThrAlaLeuIle 90
             ::: ||||||
Db   351 CTGCCCCGGGCCAATGAATTGGGTCGTGAGTAGTGATAGCAGATACAGCCCTCATC 410

Qy   78   GluAspValAlaValAlaValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSer 97
             :::
Db   201 GGGCGACACGGCCGAGATCCAAGGGACACGGCCCTAATCACGCGCCCTCGAGCCCCGCT 260

Qy   91   ThrArgProGlyAlaProSerArgArgLysGlyGluValAlaAspMetLysMetLysGluValAlaLeuGlu 110
             ||||| ::: |||
Db   411 ACCCGTCCATGATACCTGCACGAACGAAGAGAGACTGAAGGCCTGCAAAAAATCTTTGAG 470

Qy   98   ArgArgLysGlyGluValAspMetLysLysGluValAlaLeuGlnLeuLeuAsnIleVal 117
             ::: |||||
Db   261 CGTAGGGCCAGAGGTGCGATGGAGTCCAAGACCTGGGCTTGGGGGTTCGAATTGTG 320

Qy   111 LysLeuGlnLeuAsnIleValGluMetLysLysAspGluAsnAlaThrLeuAspGlyValGlyGlyAsp 130
             ::|| ||||| |||
Db   471 GAGCTGAAATTCCGAGTCTGCGAACTCAGTGAACTCTGGATGCAAGTGAT 530

Qy   118 GluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspCysGlyAspGlyAspCysGlyAspCysCysPheGlyAspArgArgGlu 137
             |||||| :::||||||
Db   321 GAAATAGGAGACAGAACGCGACCTGGATGGCCACTGACGTTCTTCACCGGCCGGGAG 380

Qy   131 ValLeuPheThrGlyArgGlnPhePheValGlyLeuSerLysSerArgThrAsnGlnArgGly 150
             |||||| ||||| :::
Db   531 ATACTTTTCACAGGTTCAGAGATTTTTGTAGGCTTGCTAAATGACCAATCTTAGAGGT 590

Qy   138 PhePheValGlyLeuSerLysSerArgThrAsnGlnArgGlyValAlaGluIleLeuAlaAspThr 157
             ::: |||||
Db   381 TTTTTCGTAGGCCTCTCCAAATGACCAATCACCGAGGAGCTGAGGATCGTGCGCACACG 440

Qy   151 AlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrValProValAlaAsp 170
             |||||| :::||

Qy   158 PheLysAspTyrAlaValSerThrValPro 167
               |||:::|||||||||||||||||||||||||
Db   441 TTCCGGGACTTCGCCGTCTCACACTGTGCCA 470

RESULT 9
US-09-918-995-14871
; Sequence 14871, Application US/09918995
; Publication No. US2003007362 3A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14871
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(452)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-14871

Alignment Scores:
Pred. No.:   4.7e-40   Length: 452
Score: 382.00   Matches: 81
Percent Similarity: 74.81%   Conservative: 20
Best Local Similarity: 60.00%   Mismatches: 34
Query Match: 26.27%   Indels: 0
DB: Gaps: 0

US-09-889-733B-2 (1-285) x US-09-918-995-14871 (1-452)

Qy   35 GlyGluGluValAspValAlaArgAlaArgGlnHisGlnLeuTyrValGlyValLeu 54
               |||||||:::: |||||||||||||||||||||||||::||
Db   39 GGCGGAGCAAACNNCNNNNAAAATNGNAAAGGGAGCACGGGGNAAACTG 98

Qy   55 GlySerLysLeuGlyLeuGlnValValGluLeuProAlaAspGluSerLeuProAspCys 74
               |||||||::|||||||||||:::||||||||||||||||||||||||||
Db   99 TGGCAACGACTGTTGCTACAGCTGCTACAGCTGCCACCTGAGAGTCATTGCCGCTGGGA 158

Qy   75 ValPheValGluAspValAlaValAlaValCysGluGluThrAlaLeuIleThrArgProGly 94
               ::: ||||||::::   :::||||||||||||||||
Db   159 CCGCTGCTTGGCGACACGGCCGTGATCCAAGGGACACGGCCCTGG 218

Qy   95 AlaProSerArgArgLysAspGluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeu 114
               :::::||:||||||||||||||||||||:|||
Db   219 AGCCCCGCTCGTAGGCCAGAGGTCCGATGGAGTCCCAAGACCCTGGGGCTC 278

Qy   115 AsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThr 134
               |||||::::|||||||||||||||||||||||||||||:||
Db   279 CGAATTGTGAAAATAGGAGACGAGAACGCGCTGGCACTGACGTTCTCTTCACC 338

Qy   135 GlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeu 154
               |||||||||||||||||||||||||||||:::
Db   339 GGCCGGGAGTTTTTCGTAGGCCTCTCCAAATGGACCAATCACCGAGAGCTG 398

Qy   155 AlaAspThrPheLysAspTyrAlaValSerThrValProValAla 169
               ::: :::|||:::|||||||||||||||||
Db   399 GCGGACACGTTCGGGACTTCGCCGTCTCCACTGTGCCAGTCTCG 443

RESULT 10
US-09-918-995-19254
; Sequence 19254, Application US/09918995
; Publication No. US2003007362 3A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19254
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-19254

Alignment Scores:
Pred. No.:   6.67e-35   Length: 403
Score: 342.00   Matches: 75
Percent Similarity: 71.32%   Conservative: 22
Best Local Similarity: 55.15%   Mismatches: 33
Query Match: 23.52%   Indels: 6
DB: Gaps: 1

US-09-889-733B-2 (1-285) x US-09-918-995-19254 (1-403)

Qy   31 ArgSerAlaLysGlyGluValAlaAspValAlaArgAlaArgGlnHisGlnLeuTyr 50
               |||||| ::::
Db   12 CGCTCTCGGATCTGGCCAA-----------------------GCTCAAAGGGAGCACGGGGTGCTG 53

Qy   51 ValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuProAlaAspGluSer 70
               |||||| :::|||||||||||||||||||:::||||
Db   54 GGAGGTAAACTGAGGCAACGACTGGGGCTACAGCTGCTAGAACTGCCACCTGAGGAGTCA 113

Qy   71 LeuProAspCysValPheValGluAspValAlaValAlaValCysGluGluThrAlaLeuIle 90
               |||:::   :::   :||||:||
Db   114 TTGCCGCTGGGACCGCTGCTTGGCGACACGGCCGTGATCCAAGGGGACACGGACCTAATC 173

Qy   91 ThrArgProGlyAlaProSerArgArgLysAspGluValAlaAspMetMetLysGluAlaLeuGlu 110
               |||:::|||:::|||||||||| :::::||||||||:::
Db   174 ACGCGGCCCTGGAGCCCCGCTCGTAGGCCACAGGTCGATGGAGTCCCAAAGCCCTGCAA 233

Qy   111 LysLeuGlnLeuAsnIleValGluMetLysGluAsnAlaThrLeuAspGlyGlyAsp 130
               |||||:::||||||||||| ||||||
Db   234 GACCTGGGGCTCCGAATTGTGAAATAGGAGACGAGAACGCGACGCCTGGATGCACTGAC 293

Qy   131 ValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGly 150
               |||||| :::|||:::||||||
Db   294 GTTCTGTTCACCGGCCGGGAGATTTTCAGAGGCCTCTCCTAATGGACCAATCACCGAGGA 353

Qy   151 AlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrVal 166
               ||| :::|||:::|||:::
Db   354 GCTGAGATCGTGCGGGACACGTTCCGGGACTTCGCCGTCTTCACTGTG 401

RESULT 11
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02

```
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
;   LENGTH: 9025608
;   TYPE: DNA
;   ORGANISM: Streptomyces avermitilis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4187715)
;   OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Alignment Scores:
Pred. No.:                          2.23e-20       Length:         9025608
Score:                              282.50         Matches:             97
Percent Similarity:                 47.65%         Conservative:        45
Best Local Similarity:              32.55%         Mismatches:         117
Query Match:                        19.43%         Indels:              39
DB:                                                Gaps:                11

US-09-889-733B-2 (1-285) x US-10-156-761-1 (1-9025608)

Qy        1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArg------------------ 12
              :::  |||||||| |||||||| ::: :::          |||
Db  2068758 TTGTGCGGCGGGGGTCCCGTGCATGATGGGCGGCCGTACCCCCTGCCGAGACGGAGACCCG 2068817

Qy       13 ---------AlaThrHisAlaValAlaArg-----------AlaLeuProGluSerLeu    25
                      ::: |||||||| :::                    ||||||||||||||
Db  2068818 CACGTGCCAGCAGAAGAGCCCTCATCCGCGCGCCCCAGCCCCCGCGCTTCGCCGAAGGGCTG 2068877

Qy       26 CysGlnHisAlaAlaLeuArgSerAlaLysGlyGluGlnValAlaArgAlaGlnGluArg  45
              :::|||||||| :::                :: |||||||||||||||||||||
Db  2068878 GTCACGCACATCGAGGCGGCG---------CAGTCGATGTCGCCCTCGCGGTCGAG 2068925

Qy       46 GlnHisGlnLeuTyrValGlySerLysLeuGlyLeuGlnValValGluLeu         65
              |||  :::              ::: ||||||| :::      ||||| :::
Db  2068926 CAGTGGAGGGCCTATGCGGAGGCGTTGCGCACGCAC---GGCTGGGAGACCGTCGAGGTG 2068982

Qy       66 ProAlaAspGluSerLeuProAspCysValPheValGluValValAlaValValCysGlu  85
              :::                   |||||||| :::              |||||
Db  2068983 GATCCGGCCGACCAGACTCGGTCGTGTTCGTCGGAGGACGCGGTGGTCATGTACCGC 2069042

Qy       86 GluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMet 105
              |||||||| :::                         |||                :::
Db  2069043 AATGTCGGCTCCGATCACCCCGGCGCCGGGCCGACTCCCGGCGCGGAGACCCGGCGTC 2069102

Qy      106 LysGluIleAlaGluGluLysGlnLeuAsnIleValGluMetLysAspGluAsnAlaThr 125
              :::|||||| ::: ::: |||       |||
Db  2069103 GAGGAGGGCCGTGCCCGGCTGCTCG---GTGAACTGGATCTGGAGACCCGGCACC 2069159

Qy      126 LeuAspGlyGlyValProPheThrGlyArgGluPheValGlyLeuSerLysArg       145
              :::              |||||| :::           :::
Db  2069160 CTCGACCGGCGGCGACGTCCTGAAGATCGGCAGACACCAGTCTAGGCGCGGGGGCGGCGGC 2069219

Qy      146 ThrAsnGlnArgGlyAlaGlyIleLeuAlaAlaAspThrPheLysAspTyrAla-----Val 163
              ||| :::                             |||
Db  2069220 ACCAACGGCGCCGGGGGTCCAGCAGCTGCGCGGGGGTCGCGGCCGCTCGAGCCCAGGTC 2069279

Qy      164 SerThrValProValAlaAspAlaGlyLeuHisLeuLysSerPheCysSerMetAlaGlyPro 183
              :::              |||||||||||||| ::: :::
Db  2069280 GTCGCCGTGCCCCGTCAGCAAGGTGCTGCACCTGAAGTCCTTCGGCTCACC---GCGCTGCCG 2069336

Qy      184 AsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGln 203
              :::                ||| :::
Db  2069337 GACGGGCACGGTCATCGGCTCACATCCCGCGTCGTCACACAGGCCGACAGGCGACTCCGCTTC 2069396

Qy      204 MetSerAspHisArgTyrAspIleLysThrLeuProAlaAspIleAlaAlaAsnCysIle 223
              :::                           |||||||||||
Db  2069397 CTCTCG---------------------------GTCCCGGAGGAGTCCCCATGTGTC 2069432

Qy      224 TyrLeuAsnIleProAsnLysGlyHisValLeuHisIleArgThrProGluGluTyrPro 243
              :::                           |||||||| |||
Db  2069433 CTGCTCGGC-----------GGATCGAAGCTGCTG-----ATGGCGGACAGCGCCCCG 2069474

Qy      244 GluSerAlaLysValTyrGluLysAspHisMetLeuIleProValSerMetSer 263
              :::::::| |||
Db  2069475 AAGACGGGACCTGCTGCCGGACCTC---GGCCACGAGACCCGTCCGTTAACATCAGC 2069531

Qy      264 GluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLys 281
              |||   |||||| |||::::::|||
Db  2069532 GAGTACGACGAGAAACTGAAGGCTGTGACATGTCTCCTGGTCCGCTGAGGGAG 2069585
```

RESULT 12
US-10-156-761-1679
; Sequence 1679, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;   APPLICANT: OMURA, SATOSHI
;   APPLICANT: IKEDA, HARUO
;   APPLICANT: ISHIKAWA, JUN
;   APPLICANT: HORIKAWA, HIROSHI
;   APPLICANT: SHIBA, TADAYOSHI
;   APPLICANT: SAKAKI, YOSHIYUKI
;   APPLICANT: HATTORI, MASAHIRA
;   TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;   FILE REFERENCE: 249-262
;   CURRENT APPLICATION NUMBER: US/10/156,761
;   CURRENT FILING DATE: 2002-05-29
;   PRIOR APPLICATION NUMBER: JP 2001-204089
;   PRIOR APPLICATION FILING DATE: 2001-05-30
;   PRIOR APPLICATION NUMBER: JP 2001-272697
;   PRIOR APPLICATION FILING DATE: 2001-08-02
;   NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1679
;   LENGTH: 774
;   TYPE: DNA
;   ORGANISM: Streptomyces avermitilis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(774)
US-10-156-761-1679

Alignment Scores:
Pred. No.:   1.73e-26   Length: 774
Score: 281.00   Matches: 91
Percent Similarity: 49.43%   Conservative: 39
Best Local Similarity: 34.60%   Mismatches: 109
Query Match: 19.33%   Indels: 24
DB: Gaps: 9

US-09-889-733B-2 (1-285) x US-10-156-761-1679 (1-774)

```
Qy       21 LeuProGluSerLeuCysGlnHisAlaAlaLeuArgSerAlaLysGlyGluValAspVal  40
              ||||| ::: |||::: |||||||||||||||||||||
Db       43 CTCGCCGAAGGGCTGTCACGCACATCGAGGCGCGG--------------CAGGTCGATGTC  90

Qy       41 AlaArgAlaGlnArgGlnHisGlnLeuTyrValGlyLysValLeuGlyLeuSerLeuGlyLeu  60
              ||| ::: |||  :::              :::
Db       91 GGCCTCGCGGTCGAGCAGTGGAGGGCTATGCGGAGGCGCTGCACGCAC---GGCTGG 147

Qy       61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluValAspVal  80
              ::: ||||| ::: :::                     |||
Db      148 GAGACCGTCGAGGTGGATCCGGCCGACGAGTCCCTGCCCGACTGCGTGTTCGTCGGAGGACGCG 207

Qy       81 AlaValValCysGluGluThrAlaLeuIleThrArgAlaLeuIleThrArgProSerArgArgLys 100
              ::: |||||||| :::              :::
Db      208 GTGGTCATGTACCGCAATGTCGGCCTGATCACCCGCGTCGCCGCCCGCACTCCCGGCCGCGGC 267

Qy      101 GluValAspMetLysMetLysGluAlaArgGlnHisHisGlnLeuGluIleValGluMetLys 120
              ||||||||||||||||  :::    ::: :::
Db      268 GAGACCCGGCGTCGAGGAGGGCCGTGCCCGGCTGCTCGGCCCCTGGCGTCTCG---GTGAACTGGATC 324

Qy      121 AspGluAsnAlaThrLeuAspGlyGlyValProPheThrGlyArgGluPheVal 140
              |||                ::: |||||||||||||||||
Db      325 TGGGAACCCGGCACCCTCGACCGGCGGCGACGTCCTGAAGATCGGCAGACACCATCTACGTA 384

Qy      141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGlyIleLeuAlaAlaAspThrPheLysAsp 160
```

```
Db  385 GGCGGGGGCGGGGGCCACCAACGGCGGGTCCAGCAGCTCGCGCGGGTCTTCGAGCCG 444
                ||||||||                    :::  |||
Qy  161 TyrAla------ValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCys 178
Db  445 CTCGGAGCCCAGGTCGTCGCCGTGCCCTCAGCAAGGTGCTGCACCTGAAGTCCTCGGTC 504
                |||||||       ||:::  ||||||||||
Qy  179 SerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeu 198
Db  505 ACC---GCGCTCCGGACGCGACGGCACGGTCATCGGTCACATCCCGCTCGTGGACAGGCCCGCG 561
                :::   ||   |||  |  ||            ::
Qy  199 LysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIle 218
Db  562 CTGTTCACTCGCTTCCTCTCG---------------GTGCCGGAGGGAGTCC 597
                :::  :::|||:|||||||                ::::::
Qy  219 AlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThr 238
Db  598 GGCTCCCATGTGGTCCCTGCTCGGC----------GGATCGAAGCTGCTG------ATG 639
                |||||||||||||||||||||:::          |||||||||||||
Qy  239 ProGluGluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIle 258
Db  640 GCGGACAGCGCCCGAGAGACGGGCGACCTGCTGCGCGACCTC--GGCCACGAGCCCGTC 696
                |||||||||||||:::::||||||||||||:::   |||||||
Qy  259 ProValSerMetSerGluLeuGluLysValAspLysGlyLeuLeuThrCysCysSerValLeu 278
Db  697 CTGGTTAACATCAGCGAGTACGAGAAACTCGAAGCGTGTGTGACATGTCTCCGGTCCGG 756
                :::||||||||||     ||:::|||  :::||||||||||
Qy  279 IleAsnLys 281
Db  757 CTGAGGGGAG 765
```

RESULT 13
US-09-783-590-5309

; Sequence 5309, Application US/09783590
; Patent No. US20020110850A1
GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2Cl
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5309
;   LENGTH: 332
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (22)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (42)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (98)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (134)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (139)

;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (232)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (257)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (263)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (270)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (272)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (312)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (319)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (326)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-5309

Alignment Scores:
Pred. No.:   Length: 332
Score: 4.26e-25   Matches: 83
Percent Similarity: 266.00   Conservative: 2
Best Local Similarity: 75.22%   Mismatches: 17
Query Match: 73.45%   Indels: 11
DB: 18.29%   Gaps: 2
                         9

US-09-889-733B-2 (1-285) x US-09-783-590-5309 (1-332)

```
Qy  130 AspValPheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArg 149
                ||||||||||||||||||||||||||||||||||
Db   12 GATGTTTATNCACAGGCAGAGAATTTTTNTGGCCTTTCCAAAAGGACAANTCAACGA 71

Qy  150 GlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrVal-ProValAl 169
                ||||||||||||||||||||||||||||||||||||||||
Db   72 GGTGCTGAAATCTTGGCTGATACTTTNAAGGACATATGCAGTGTCTCCACAGTGCCCAGTGGC 131

Qy  169 a-AspGlyLeuHisLeuLysSerPheCys-SerMetAlaGlyPro--AsnLeuIleAla 187
                | ||||||||||||||||||||||||| |||||||||||
Db  132 AGNATGGNTTGCATTTAAAGAGTTTCTGCCAGCATGGTTGGGCCTGAACCTGAATCGCA 191

Qy  188 IleGlySerSerGlu-SerAla-GlnLysAlaLeuLysIleMetGlnGlnMetSerAsp- 206
                ||||||||||    ||:::|||||||||||||||||||||||||||:::
Db  192 ATTGGGTGTAGTGAAATCTGCCACAGAAGGCCCTTAAGATNATGCACAAATGAGTGAAC 251

Qy  207 HisArg-TyrAspLysLeuThrVal-ProAsp-AspIleAlaAlaAsnCysIleTyrLeu 225
                ||||| ||||||||||||||  ||||  |||:::
Db  252 CACCGNTAAGGNCAAACTNANTGTGCCCTGATTGACATAGCAGCAGCAAATTGTATATGATCT 311

Qy  226 AsnIle--ProAsnLysGly 231
                :::  ||||||||||
Db  312 NAATTATNCCCCAANAAAGGG 332
```

RESULT 14
US-10-450-826-8/c

; Sequence 8, Application US/10450826
; Publication No. US20040101818A1
GENERAL INFORMATION:
; APPLICANT: Ji, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathon S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Eistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence

; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiati
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.1
US-10-450-826-8
; SEQ ID NO 8
; LENGTH: 83946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AL078459
US-10-450-826-8

Alignment Scores:
Pred. No.:
Score: 2.64e-21   Length: 83946
Percent Similarity: 265.50   Matches: 77
Best Local Similarity: 45.18%   Conservative: 12
Query Match: 39.09%   Mismatches: 53
DB: 18.26%   Indels: 55
                        17   Gaps: 6

US-09-889-733B-2 (1-285) x US-10-450-826-8 (1-83946)

Qy   85 GluGluThrAlaLeuIleThrArgProGly------------------------------ 94
          |||||:|||||||::||||||
Db 18988 GAGGAGGCAGCCCTGCTTACTCAGCCAAAGATATCAAAACCCCTGCTGGCCTTACATGT 18929

Qy   95 ------------AlaProSerArgArgLysGluValAspMet 104
                      ::||||
Db 18928 GGCCATTTGCATCTTACTGATAAGTCTTACGACAAAGACTAGAAAAAATAATAGACGGT 18869

Qy   105 MetLysGluAlaLeuGluLysGlnLeuAsnIleValGluMetLysAspGluAsnAla 124
          ::||||||
Db 18868 TTTAAAACCTTAAAAGAAGAGAGACAAAACACATACTTTTAAAGAGCAACAAAAAACCCT 18809

Qy   125 ThrLeuAspGlyGlyAsp-------------ValLeuPhe------------Thr 134
                              ||||||| |||||
Db 18808 TGCAATTCTTCCAAGGACTCATAGTTAATCAAGTTTCCTTTTCTGGTTCTAAGTGACT 18749

Qy   135 GlyArg---------GluPheValGlyLeuSerLysArgThrAsnGlnArgGlyAla 151
          |||||| |||||| ::::||
Db 18748 GGAGGTAATAGATGGGAATT------------------GAACAGAGAAACAGCCTAAAGTGT 18704

Qy   152 GluIleLeuAlaAspThrPheLysAspThrAlaValSerThrValProValAlaAspGly 171
          ::|| :::
Db 18703 CAGTTCCAAGCTTGTATCCTACCACCACTGCCTTCCTTCTGTTTGAAAGTGTGGGAGGTTA 18644

Qy   172 LeuHisLysLysSerPheCysSerMetAlaGlyProAsnLeuIleLeuIleGlySerSer 191
          :::|||
Db 18643 CTCTTCTTAGGCCTGCTTGTTTC---------------------------- 18620

Qy   192 GluSerAlaGlnLysLysIleMetGlnMetSerAspHisArgTyrArgAspTyrAspLys 211
          :::||
Db 18619 --------------------TCTTTATAGATCATGCAACAGATGAGTGACCACCCTACGACAAA 18575

Qy   212 LeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGly 231
          |||||||
Db 18574 CTCACTGTGCCTGATGATGCATAGCAACATGACGAAACTGTATATATCCCAACAAAGGG 18515

Qy   232 HisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysVal 248
          ||||||
Db 18514 CACGTCTTGCTGCTGAACCCCGGAGAGATCCAGAAAGTGCAAAGGTA 18464

RESULT 15
US-09-736-457-404/c
; Sequence 404, Application US/09736457
; Patent No. US2002016837A1
; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-404

Alignment Scores:
Pred. No.:
Score: 5.28e-21   Length: 239
Percent Similarity: 233.00   Matches: 49
Best Local Similarity: 78.48%   Conservative: 13
Query Match: 62.03%   Mismatches: 17
DB: 16.02%   Indels: 0
                        9   Gaps: 0

US-09-889-733B-2 (1-285) x US-09-736-457-404 (1-239)

Qy   62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
          :::::|||||||||||||| :::|||||||||||
Db   239 CTGCTAGAACTGCCACCTGAGGAGTCATTGCCGCTGGGACCGCTGCTTGGCCGACACGGCC 180

Qy   82 ValValCysGluGlyThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
          ||||:::   ::|||||||||||
Db   179 GTGATCCCAAGGGACACCGGCGCCTAATCACGCCCCGCTTCGTCGGTAGGCCAGAG 120

Qy   102 ValAspMetMetLysGluAlaLeuGluLysGlnLeuAsnIleValGluMetLysAsp 121
          ||||:::   :::|||
Db   119 GTCGATGGAGTCCCGCAAGACCCTGCAAGACCTGGGGCTCCGAATTGTGGAAATAGGAGAC 60

Qy   122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
          :::|| |||
Db   59 GAGAACGGCAGCGCTGGATGGCACTGAGCGTTCTTCACCGGCCGGGAGTTTTTCGTA 3

Search completed: June 8, 2004, 02:56:42
Job time : 6580 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 00:22:57 ; Search time 3209 Seconds
                         (without alignments)
                         3849.412 Million cell updates/sec

Title: US-09-889-733B-2
Perfect score: 285
Sequence: 1 MAGLGHPSAFGRATHAVVRA.........EKVDGLLTCCSVLINKKVDS 285

Scoring table: OLIGO
                Xgapop 60.0 , Xgapext 60.0
                Ygapop 60.0 , Ygapext 60.0
                Fgapop   6.0 , Fgapext   7.0
                Delop   6.0 , Delext   7.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934735

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=x1p
 -MODEL=frame+_p2n.model -DEV=x1p
 -Q=/cgn2_1/USPTO_spool_p/US09889733/runat_07062004_083557_8353/app_query.fasta_1.455
 -DB=GenEmbl -QFMT=fastap -SUFFIX=olip2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09889733 @CGN_1_1_5600 @runat_07062004_083557_8353 -NCPU=6 -ICPU=3
 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :   GenEmbl:*
          1: gb_ba:*
          2: gb_htg:*
          3: gb_in:*
          4: gb_om:*
          5: gb_ov:*
          6: gb_pat:*
          7: gb_ph:*
          8: gb_pl:*
          9: gb_pr:*
         10: gb_ro:*
         11: gb_sts:*
         12: gb_sy:*
         13: gb_un:*
         14: gb_vi:*
         15: em_ba:*
         16: em_fun:*
         17: em_hum:*
         18: em_in:*
         19: em_mu:*
         20: em_om:*
         21: em_or:*
         22: em_ov:*
         23: em_pat:*
         24: em_ph:*
         25: em_pl:*
         26: em_ro:*
         27: em_sts:*
         28: em_un:*
         29: em_vi:*
         30: em_htg_hum:*
         31: em_htg_inv:*
         32: em_htg_other:*
         33: em_htg_mus:*
         34: em_htg_pln:*
         35: em_htg_rod:*
         36: em_htg_mam:*
         37: em_htg_vrt:*
         38: em_sy:*
         39: em_htgo_hum:*
         40: em_htgo_mus:*
         41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	285	100.0	858	6	BD251986	BD251986 Screen me
2	285	100.0	858	6	AX032835	AX032835 Sequence
3	259	90.9	1624	9	BC033680	BC033680 Homo sapi
4	259	90.9	1633	6	BD094073	BD094073 Shear str
5	259	90.9	1633	6	AB001915	AB001915 Homo sapi
6	259	90.9	4208	6	AX548046	AX548046 Sequence
7	246	86.3	3737	9	BC043235	BC043235 Homo sapi
8	184	64.6	3908	9	HSM808292	BX648145 Homo sapi
9	106	37.2	1120	10	BC034505	BC034505 Mus muscu
10	106	37.2	3008	9	AX683109	AX683109 Sequence
11	106	37.2	3008	10	D86041	D86041 Rattus norv
12	84	29.5	2895	9	AK098405	AK098405 Homo sapi
13	76	26.7	104071	2	AL162260	AL162260 Homo sapi
14	76	26.7	158405	2	AC019261	AC019261 Homo sapi
15	76	26.7	168953	2	AL360219	AL360219 Human DNA
16	69	24.2	2289	10	AK128909	AK128909 Mus muscu
17	49	17.2	83946	9	HSA123M24	AL078459 Human DNA
18	48	16.8	1344	9	BC056074	BC056074 Xenopus l
19	46	16.1	3420	10	BC060266	BC060266 Mus muscu
20	38	13.3	186434	2	AC126412	AC126412 Mus muscu
21	38	13.3	228940	2	AC126890	AC126890 Rattus no
22	34	11.9	186434	2	AC126412	AC126412 Mus muscu
23	34	11.9	200007	2	AC123684	AC123684 Mus muscu
24	34	11.9	212111	2	AC094482	AC094482 Rattus no
25	31	10.9	1150	5	AB095027	AB095027 Gallus ga
26	27	9.5	420	6	AX898477	AX898477 Sequence
27	27	9.5	420	6	BD034010	BD034010 Sequence
28	26	9.1	110000	2	AC144263_2	Continuation (3 of
29	26	9.1	304208	2	AC143028_2	AC143028 Macaca mu
30	25	8.8	626	11	BV064692	BV064692 S212P6014
31	19	6.7	58539	2	AC137903	AC137903 Mus muscu
32	16	5.6	1377	3	AY060959	AY060959 Drosophil
33	15	5.3	858	6	BD251987	BD251987 Screen me
34	15	5.3	858	6	AX032837	AX032837 Sequence
35	15	5.3	1228	6	AX780306	AX780306 Sequence
36	15	5.3	1263	9	BC001435	BC001435 Homo sapi
37	15	5.3	1270	10	AF004106	AF004106 Mus muscu
38	15	5.3	1310	10	BC003328	BC003328 Mus muscu
39	15	5.3	1342	9	AK098671	AK098671 Homo sapi
40	15	5.3	1351	6	AX779950	AX779950 Sequence
41	15	5.3	1351	9	AF087894	AF087894 Homo sapi
42	15	5.3	1376	9	AX011712	AX011712 Homo sapi
43	15	5.3	1376	6	BD226323	BD226323 Pancreati
44	15	5.3	1397	6	AR379492	AR379492 Sequence
45	15	5.3	1709	9	AK026191	AK026191 Homo sapi

ALIGNMENTS

RESULT 1

RESULT 1

LOCUS   BD251986   858 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION Screen method.
ACCESSION   BD251986
VERSION   BD251986.1 GI:33061756
KEYWORDS   JP 2002535001-A/1.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 858)
  AUTHORS   Vallance,P.J.T., Leiper,J.M., Whitley,G.S.J. and Charles,I.G.
  TITLE   Screen method
  JOURNAL   Patent: JP 2002535001-A 1 22-OCT-2002;
            UNIVERSITY COLLEGE LONDON
COMMENT   OS   Homo sapiens (human)
            PN   JP 2002535001-A/1
            PD   22-OCT-2002
            PF   26-JAN-2000 JP 2000596131
            PR   26-JAN-1999 GB   9901705.5,04-JUN-1999 GB   9913066.8 PI
            PATRICK JOHN THOMPSON VALLANCE,JAMES MITCHELL LEIPER,GUY ST   PI
            JOHN WHITLEY,
            PI   IAN GEORGE CHARLES
            PC   C12N15/09,A01K67/027,A61K31/198,A61K38/00,A61K45/00,A61P3/06,
            PC   A61P9/02,
            PC   A61P9/04,A61P9/10,A61P9/12,A61P13/12,A61P25/06,A61P25/18, PC
            A61P25/28,
            PC   A61P29/00,A61P31/04,A61P35/00,C07K16/40,C12N1/15,C12N1/19, PC
            C12N1/21,
            PC   C12N5/10,C12N9/78,C12Q1/02,C12Q1/34,G01N33/15,G01N33/50 PC
            ,C12N15/00,C12N5/00,
            CC   A61K37/02
            CC   Screen method
FH   Key   Location/Qualifiers
FT
FEATURES   Location/Qualifiers
     source   1..858
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:   Length: 858
Score: 1.25e-306   Matches: 285
Percent Similarity: 285.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 6

US-09-889-733B-2 (1-285) x BD251986 (1-858)

Qy   1 MetAlaGlyLeuGlyHisProSerLeuArgAlaPheGlyArgAlaThrHisAlaValValArgAla 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGCCGGCCTCGGCCACCCTTCGGCCCGGGCCTTCGGCCGGGCCACCCACGCGTGTGCGGGCG 60

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTACCCGAGTCGCCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 120

Qy   41 AlaArgAlaGlnArgGlnHisGlnLeuTyrValGlyValSerGlyLysLeuGlyLeu 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCCCGCGCGCAGCGCCAGCACCAGCTCTACGTGGGCGTGTCGGGCAGCAAGCTGGGGCTG 180

Qy   61 GlnValGluLeuProAlaAspCysValSerProLysProValPheGluValGluAspVal 80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAGGTGGTGGAGCTGCCCGCCGACGAGCGCCCTTCCGGAACGTGCCTCTTCGTGGAGGACGTG 240

Qy   81 AlaValValCysGluGlyThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GCCGTGGTGTGCGAGGAGGACGGGCCCTCATCACCCGACCCGGGCGCGCCAGCCGGAGCGGAAG 300

Qy   101 GluValAspMetMetLysGluAlaLeuGluAlaGluGlnLysLeuGlnLeuAsnIleValGluMetLys 120

RESULT 3

```
LOCUS       BC033680                1624 bp    mRNA    linear   PRI 07-OCT-2003
DEFINITION  Homo sapiens dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
            clone MGC:45161 IMAGE:5189970), complete cds.
ACCESSION   BC033680
VERSION     BC033680.1  GI:21707414
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1624)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   MEDLINE  22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1624)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Lewis,L.P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 68 Row: m Column: 1
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 6912327.
FEATURES             Location/Qualifiers
     source          1..1624
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

ORIGIN

VLINKKVDS"

Alignment Scores:		
Pred. No.:		
Score:	1.25e-306	
Percent Similarity:	285.00	
Best Local Similarity:	100.00%	
Query Match:	100.00%	
DB:	100.00%	

	Length:	858
	Matches:	285
	Conservative:	0
	Mismatches:	0
	Indels:	0
	Gaps:	0

US-09-889-733B-2 (1-285) x AX032835 (1-858)

```
Qy    1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db    1 ATGGCCGGGCTCGGCCACCCTCGGCCTTCGGCCGGGCCACCCACGCGCGTGGTGCGGGCG 60

Qy   21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAlaAspVal 40
Db   61 CTACCCGAGTCGCTCTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC 120

Qy   41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLysGlyLeu 60
Db  121 GCCCGCGCCGGAACGGGCAGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTG 180

Qy   61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db  181 CAGGTGGTGGAGCTGCCCGCCGACGAGAGCCCTTCCGGACTGCCTCTTCGTGGAGGACGTG 240

Qy   81 AlaAlaValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
Db  241 GCCGTGGTGTGCGAGGAGACGGCGCTCATCACCCGACCCGGGGCGCCGAGCCCGAGGAAG 300

Qy  101 GluValAspMetMetLysGluLeuGlyLeuGlnLeuAsnIleValGluMetLys 120
Db  301 GAGGTGGACATGATGAAAGAGCATTAGAAAACTTCAGCTCAATATAGTAGATGATGAA 360

Qy  121 AspGluAsnAlaThrLeuAspThrGlyAspGlyValLeuPheThrGlyArgGluPheVal 140
Db  361 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTG 420

Qy  141 GlyLeuSerLysArgThrArgGlnInArgGlyValGlyAlaGluIleLeuAlaLeuLysAspAsp 160
Db  421 GGCCTTTCCAAAAGGACAAGATCAACGAGGTGCTGAAATCTTGGCTGAAACTTTTAAGGAC 480

Qy  161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db  481 TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATG 540

Qy  181 AlaGlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnInLysAlaLeuLysIle 200
Db  541 GCTGGGCCTAACCTGATCGCAATTGGCTGGGTCTTAGTGAATCTGCAACAGAAGGCCCTTAAGATC 600

Qy  201 MetGlnGlnMetSerAspHisArgTyrAspProLeuThrValProAspAspIleAlaAla 220
Db  601 ATGCAACAGATGAGTGACCACCGCTCACGACAAACTCACTGTGCCTGATGACATAGCAGCA 660

Qy  221 AsnCysIleTyrLeuAsnIleProAsnLysLysGlyHisValLeuValHisArgThrProGlu 240
Db  661 AACTGTATATATCTAAATATCCCAACAAAGGGCACGTTCTTGTCTGCACCGACCCCCGGAA 720

Qy  241 GluTyrProGluSerAlaLysValArgGluLysLeuLysAspHisMetLeuIleProVal 260
Db  721 GAGTATCCAGAAAGTGCAAAGGTTAGAGAAAAACTGAAGGACCATATGCTGATCCCCGTG 780

Qy  261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysSerValLeuIleAsn 280
Db  781 AGCATGTCTGAACTGGAAAAAGGTGGATGGGCTGCTCACCTGCTCTGCTCAGTTTTAATTAC 840

Qy  281 LysLysValAspSer 285
Db  841 AAGAAGGTAGACTCC 855
```

```
          /clone="MGC:45161 IMAGE:5189970"
          /tissue_type="Colon, Kidney, Stomach, adult, whole pooled"
          /clone_lib="NIH_MGC_116"
          /lab_host="DH10B"
gene      /note="Vector: pCMV-SPORT6"
          1..1624
          /gene="DDAH1"
CDS       /note="synonym: DDAH"
          /db_xref="LocusID:23576"
          /db_xref="MIM:604743"
          100..957
          /codon_start=1
          /product="DDAH1 protein"
          /protein_id="AAH33680.1"
          /db_xref="GI:21707415"
          /db_xref="LocusID:23576"
          /translation="MAGLGHPAAFGRATHAVVRALPESLGQHALRSAKGEEVDVARAE
          RQHQLIVGVLGSKLGLQVVELPADSLPDCVFVEDVAVVCEETALITRPGAPSRRKEV
          DMMKEALEKLQLNIVEMKDENATLDGGDVLFTGREFFVGLSKRTNQRGAEILADTFKD
          YAVSTVPVADGLHLHKSFCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDI
          AANCIVLNIPNKGHVLLHRTPEEYPBSAKVYEKLKDHMLIPVSMSBLEKVDGLLTCCS
          VLINKKVDS"
misc_feature  133..936
          /note="Amidinotransf; Region: Amidinotransferase. This
          family contains glycine (EC:2.1.4.1) and inosamine
          (EC:2.1.4.2) amidinotransferases, enzymes involved in
          creatine and streptomycin biosynthesis respectively. This
          family also includes arginine deiminases, EC:3.5.3.6.
          These enzymes catalyse the reaction: arginine + H2O <=>
          citrulline + NH3. Also found in this family is the
          Streptococcus anti tumor glycoprotein"
          /db_xref="CDD:pfam02274"
```

ORIGIN

Alignment Scores:
```
Pred. No.:            1.83e-277   Length:           1624
Score:               259.00      Matches:          259
Percent Similarity:  100.00%     Conservative:     0
Best Local Similarity: 100.00%   Mismatches:       0
Query Match:         90.88%      Indels:           0
DB:                              Gaps:             9
```

US-09-889-733B-2 (1-285) x BC033680 (1-1624)

```
Qy   27 GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln 46
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  178 CAGCAGCGCCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTCGCCCGCGCCGAACGGCAG 237

Qy   47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  238 CACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCG 297

Qy   67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValAlaValCysGluGlu 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  298 GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAG 357

Qy   87 ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLys 106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  358 ACGGCCCTATCACCGACCCGGGGCCCGAGCCCGGAGGAAGGAGGTTGACATGATGAAA 417

Qy  107 GluValAspValGlyLysGluAsnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu 126
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  418 GAAGCAATTAGAAAAACTTCAGCTCAATATAGTAGATGAGATGAAAATCAACTTTA 477

Qy  127 AspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThr 146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  478 GATGGCGGAGATGTTTTATTCACAGGCAGAGAGAATTTTTGTGGGCCTTTCCAAAAGGACA 537

Qy  147 AsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrVal 166
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  538 AATCAACGAGGTGCTGAAATCTTCGCTGATACTTTAAGGACTATGCAGTCTCCACAGTG 597

Qy  167 ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle 186
```

(right column, continued alignments)

```
Db  598 CCAGTTGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGGTTGGCCTAACCTGATC 657

Qy  187 AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp 206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  658 GCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGAC 717

Qy  207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleValLeuAsn 226
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  718 CACCGGTACGACCAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAAT 777

Qy  227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla 246
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  778 ATCCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAAGAGTATCCAGAAAGTGCA 837

Qy  247 LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  838 AAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAA 897

Qy  267 LysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  898 AAGGTCGATGGGCTGCTCACCTGCTGCTCAGTTTAATTAACAACAAGTAGACTCC 954
```

RESULT 4
BD094073
```
LOCUS       BD094073    1633 bp    DNA    linear    PAT 27-AUG-2002
DEFINITION  Shear stress-responsive DNAs.
ACCESSION   BD094073
VERSION     BD094073.1   GI:22639661
KEYWORDS    WO 0125427-A/34.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1633)
  AUTHORS   Nojima,H., Yoshisue,H., Obayashi,M., Ota,T., Kawabata,A.,
            Sakurada,K., Kuga,T., Sekine,S., Nakamura,Y. and Sugan,S.
  TITLE     Shear stress-responsive DNAs
  JOURNAL   Patent: WO 0125427-A 34 12-APR-2001;
            KYOWA HAKKO KOGYO CO LTD,HIROSHI NOJIMA,HAJIME YOSHISUE, MASAYA
            OBAYASHI, TOSHIO OTA,AYAKO KAWABATA,KAZUHIRO SAKURADA,TETSURO KUGA,
            SUSUMU SEKINE, YUSUKE NAKAMURA,SUMIO SUGANO
COMMENT     OS  Homo sapiens (human)
            PN  WO 0125427-A/34
            PD  12-APR-2001
            PF  02-OCT-2000 WO 2000JP006840
            PR  01-OCT-1999 JP 99P 280976
            PI  HIROSHI NOJIMA, HAJIME YOSHISUE,MASAYA OBAYASHI, TOSHIO OTA, PI
            AYAKO KAWABATA,
            PI  KAZUHIRO SAKURADA,TETSURO KUGA, SUSUMU SEKINE, YUSUKE NAKAMURA,
            PI  SUMIO SUGANO
            PC  C12N15/12,C07K14/435,C07K16/18,C12P21/02,C12Q1/68,A61K38/00,
            PC  A61K39/395,
            PC  A61K48/00,A61P9/10,G01N33/50,G01N33/53
            CC
FH          Key            Location/Qualifiers
FT          CDS            (323)..(1177)
```
FEATURES   Location/Qualifiers
```
source          1..1633
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
```
ORIGIN

Alignment Scores:
```
Pred. No.:           1.84e-277   Length:           1633
Score:               259.00      Matches:          259
Percent Similarity:  100.00%     Conservative:     0
Best Local Similarity: 100.00%   Mismatches:       0
Query Match:         90.88%      Indels:           0
DB:                              Gaps:             6
```

US-09-889-733B-2 (1-285) x BD094073 (1-1633)

```
REFERENCE    2  (bases 1 to 1633)
AUTHORS      Kimoto,M.
TITLE        Direct Submission
JOURNAL      Submitted (15-MAR-1997) Masumi Kimoto, Okayama Prefectural
             University, Faculty of Health and Welfare Science, Department of
             Nutritional Science; Kuboki 111, Soja, Okayama 719-1197, Japan
             (E-mail:kimoto@fhw.oka.-pu.ac.jp, Tel:0866-94-2158,
             Fax:0866-94-2158)
FEATURES             Location/Qualifiers
     source          1..1633
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="Kidney"
     CDS             323..1180
                     /EC_number="3.5.3.18"
                     /codon_start=1
                     /product="NG,NG-dimethylarginine dimethylaminohydrolase"
                     /protein_id="BAA37117.1"
                     /db_xref="GI:4160666"
                     /translation="MAGLGHPAAFGRATHAVVRALPESLGQHALRSAKGEVDVARAE
                     RQHQLYVGLVGSKLGLQVVELPADESLPDCVFVEDVAVVCEETALITRPGAPSRRKEV
                     DMKRALEKLQLNIVEMKDENATLDGGDVLFTGREFFVGLSKRTNQRGAELLADTFKD
                     YAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDI
                     AANCIYLNIPNKGHVLLHRTPEEYPESAKVYEKLKDHMLIPVSMSLEKVDGLLTCCS
                     VLINKKVDS"

ORIGIN

Alignment Scores:
Pred. No.:                          Length:           1633
Score:                   1.84e-277  Matches:          259
Percent Similarity:      259.00     Conservative:     0
Best Local Similarity:   100.00%    Mismatches:       0
Query Match:             100.00%    Indels:           0
DB:                      90.88%     Gaps:             0
                         9

US-09-889-733B-2 (1-285) x AB001915 (1-1633)

Qy    27  GlnHisAlaLeuArgSerAlaLysGlyGluGluValAspValAlaArgAlaAlaGluArgGln  46
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   401  CAGCACGGCCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTGGCCCGCGGGAACGGCGCAG  460

Qy    47  HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro  66
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   461  CACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCCTGCAGGTGGTGGAGCTGCCG  520

Qy    67  AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu  86
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   521  GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAG  580

Qy    87  ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLys  106
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   581  ACGGCCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAA  640

Qy   107  GluAlaLeuGluLysGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu  126
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   641  GAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTA  700

Qy   127  AspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThr  146
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   701  GATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTGGGCCTTTCCAAAAGGACA  760

Qy   147  AsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrVal  166
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   761  AATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTG  820

Qy   167  ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle  186
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   821  CCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCCTAACCTGATC  880

Qy   187  AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp  206
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   881  GCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGAC  940
```

(left column, rotated)
```
Qy    27  GlnHisAlaLeuArgSerAlaLysGlyGluGluValAspValAlaArgAlaAlaGluArgGln  46
Db   401  CAGCACGGCCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTGGCCCGCGGGAACGGCGCAG  460

Qy    47  HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro  66
Db   461  CACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCCTGCAGGTGGTGGAGCTGCCG  520

Qy    67  AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu  86
Db   521  GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAG  580

Qy    87  ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLys  106
Db   581  ACGGCCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAA  640

Qy   107  GluAlaLeuGluLysGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu  126
Db   641  GAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTA  700

Qy   127  AspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThr  146
Db   701  GATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTGGGCCTTTCCAAAAGGACA  760

Qy   147  AsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrVal  166
Db   761  AATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTG  820

Qy   167  ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle  186
Db   821  CCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCCTAACCTGATC  880

Qy   187  AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp  206
Db   881  GCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGAC  940

Qy   207  HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn  226
Db   941  CACCGTTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAAT  1000

Qy   227  IleProAsnLysGlyHisValLeuLeuHisArgThrProGluTyrProGluSerAla  246
Db  1001  ATCCCCAACAAAGGCACGTTCTTGCTGCACAGAGTATCCAGAAAGTGCA  1060

Qy   247  LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu  266
Db  1061  AAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAA  1120

Qy   267  LysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer  285
Db  1121  AAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATTAACAAGAAAGTAGACTCC  1177

RESULT 5
AB001915
LOCUS        AB001915          1633 bp    mRNA     linear      PRI 21-SEP-2000
DEFINITION   Homo sapiens mRNA for NG,NG-dimethylarginine
             dimethylaminohydrolase, complete cds.
ACCESSION    AB001915
VERSION      AB001915.1  GI:4160665
KEYWORDS     NG,NG-dimethylarginine dimethylaminohydrolase.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (sites)
AUTHORS      Kimoto,M., Miyatake,S., Sasagawa,T., Yamashita,H., Okita,M.,
             Oka,T., Ogawa,T. and Tsuji,H.
TITLE        Purification, cDNA cloning and expression of human
             NG,NG-dimethylarginine dimethylaminohydrolase
JOURNAL      Eur. J Biochem. 258 (2), 863-868 (1998)
MEDLINE      99089653
PUBMED       9874257
```

```
Qy  127 AspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThr 146
Db  746 GATGGGCGGAGATGTTTTTATTCACAGGCAGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACA 805

Qy  147 AsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrVal 166
Db  806 AATCAACGAGGTGCGAAATCTTGCTGCTGATACTTTAAGGACTATGCAGTGTCCCAGTG 865

Qy  167 ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle 186
Db  866 CCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGGCCTAACCTGATC 925

Qy  187 AlaIleGlySerGluSerArgAlaGlnLysAlaLeuLysIleMetGlnMetSerAsp 206
Db  926 GCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTAAGATCATGCAACAGATGAGTGAC 985

Qy  207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAlaAsnCysIleTyrLeuAsn 226
Db  986 CACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTTAAAT 1045

Qy  227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla 246
Db 1046 ATCCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAAGAGTATCCAGAAAGTGCA 1105

Qy  247 LysValTyrGluLysLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
Db 1106 AAGGTTTATGAAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAA 1165

Qy  267 LysValAspGlyLeuLeuThrCysCysSerValLeuIleLeuAsnLysLysValAspSer 285
Db 1166 AAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTAATTAACAAGAAGTAGACTCC 1222
```

RESULT 7
BC043235
LOCUS   BC043235   3737 bp   mRNA   linear   PRI 30-SEP-2003
DEFINITION Homo sapiens dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
            clone IMAGE:5295723), partial cds.
ACCESSION   BC043235
VERSION   BC043235.2   GI:34783628
KEYWORDS   .
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 3737)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22388257
PUBMED   12477932
REFERENCE   2 (bases 1 to 3737)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (09-JAN-2003) National Institutes of Health, Mammalian

---

```
Qy  207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAlaAsnCysIleTyrLeuAsn 226
Db  941 CACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTTAAAT 1000

Qy  227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla 246
Db 1001 ATCCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAAGAGTATCCAGAAAGTGCA 1060

Qy  247 LysValTyrGluLysLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
Db 1061 AAGGTTTATGAAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAA 1120

Qy  267 LysValAspGlyLeuLeuThrCysCysSerValLeuIleLeuAsnLysLysValAspSer 285
Db 1121 AAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTAATTAACAAGAAGTAGACTCC 1177
```

RESULT 6
AX548046
LOCUS   AX548046   4208 bp   DNA   linear   PAT 26-NOV-2002
DEFINITION Sequence 22 from Patent WO02066654.
ACCESSION   AX548046
VERSION   AX548046.1   GI:25813142
KEYWORDS   .
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Astromoff,A., Au-Young,J., Baughn,M.R., Ding,L., Duggan,B.M.,
            Forsythe,I.J., Gietzen,K.J., Griffin,J.A., Lee,E.A., Lu,Y.,
            Richardson,T.W., Ring,H.Z., Sanjanwala,M.M., Swarnakar,A.,
            Walla,N.K., Warren,B.A., Xu,Y., Yue,H. and Zebarjadian,Y.
TITLE   Drug metabolizing enzymes
JOURNAL   Patent: WO 02066654-A 22 29-AUG-2002;
            Incyte Genomics, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..4208
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7671089CB1"

ORIGIN

Alignment Scores:
Pred. No.:
Score: 259.00   Length: 4208
Percent Similarity: 100.00%   Matches: 259
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 90.88%   Mismatches: 0
DB: Indels: 0
                        4.47e-277   Gaps: 0

US-09-889-733B-2 (1-285) x AX548046 (1-4208)

```
Qy   27 GlnHisAlaLeuArgSerAlaLysGlyGluValAspGluLeuAlaAspValAlaArgAlaArgGln 46
Db  446 CAGCACGCGCTGAGAAGCGCCAAGGGCGAGGACGTCGACGAGCTGGACGTCGCCCGGGCAAGCGCAG 505

Qy   47 HisGlnLeuTyrValGlyLeuValLeuGlySerLeuGlyLeuGlnValValGluLeuPro 66
Db  506 CACCAGCTCTACGTGGGCGTGGTGCTGGGCAGCAAGCTGGGCCTGCAGGTGGTGGAGCTGCCG 565

Qy   67 AlaAspGluSerLeuProAspCysValPheValGluValAlaValAlaValCysGluGlu 86
Db  566 GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTCGGTGTGCGAGGAG 625

Qy   87 ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysProSerArgAspMetMetLys 106
Db  626 ACGGCCCTCATCACCCGACCCGGCGCCCCGAGCCGGAGGAAGCCGAGCAGGGTTGACATGATGAAA 685

Qy  107 GluAlaLeuGluLysGlnLeuGlnIleLeuAsnIleValGluMetLysAspGlyLeuMetAlaThrLeu 126
Db  686 GAAGCATTAGAAAAACTTCAGCTCAATATAGTAGTAGAGATGAAAGATGAAAAATGCAACTTTA 745
```

REMARK
```
        Gene Collection (MGC), Cancer Genomics Office, National Cancer
        Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
        USA
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
        On Sep 16, 2003 this sequence version replaced gi:28175756.
        Contact: MGC help desk
        Email: cgapbs-r@mail.nih.gov
        Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
        cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
        Toshiyuki and Piero Carninci (RIKEN)
        cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: Genome Sequence Centre,
        BC Cancer Agency, Vancouver, BC, Canada
        info@bcgsc.bc.ca
        Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
        Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
        Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
        Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
        Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
        Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
        Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
        George Yang, Scott Zuyderduyn, Marco Marra.

        Clone distribution: MGC clone distribution information can be found
        through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
        Series: IRAK Plate: 76 Row: j Column: 15.
FEATURES             Location/Qualifiers
     source          1..3737
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5295723"
                     /tissue_type="Testis"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
     gene            <1..3737
                     /gene="DDAH1"
                     /note="synonym: DDAH"
                     /db_xref="LocusID:23576"
                     /db_xref="MIM:604743"
     CDS             <1..741
                     /gene="DDAH1"
                     /codon_start=1
                     /product="DDAH1 protein"
                     /protein_id="AAH43235.2"
                     /db_xref="GI:34783629"
                     /db_xref="LocusID:23576"
                     /db_xref="MIM:604743"
                     /translation="VARAERQHQLYVGVLGSKLGLQVVELPADESLPDCVFVEDVAVV
                     CEETALITRPGAPSRRKEVDMKSALEKLQLNIVEMKDENATLDGGDVLFTGREFFVG
                     LSKRTNQRGAEILADTFKDYAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALK
                     IMQMSDHRYDKLTVPDDIAANCIYLNIPNKGHVLLHRTPEEYPESAKVYEKLKDHML
                     IPVSMSELEKVDGLLTCCSVLINKKVDS"
     misc_feature    1..720
                     /gene="DDAH1"
                     /note="Amidinotransf; Region: Amidinotransferase. This
                     family contains glycine (EC:2.1.4.1) and inosamine
                     (EC:2.1.4.2) amidinotransferases, enzymes involved in
                     creatine and streptomycin biosynthesis respectively. This
                     family also includes arginine deiminases, EC:3.5.3.6.
                     These enzymes catalyze the reaction: arginine + H2O <=>
                     citrulline + NH3. Also found in this family is the
                     streptococcus anti tumor glycoprotein"
                     /db_xref="CDD:pfam02274"
ORIGIN
```

Alignment Scores:
```
Pred. No.:            1.14e-262    Length:        3737
Score:               246.00       Matches:        246
Percent Similarity:  100.00%      Conservative:     0
Best Local Similarity: 100.00%    Mismatches:       0
Query Match:          86.32%      Indels:           0
```

```
DB:                    9          Gaps:          0

US-09-889-733B-2 (1-285) x BC043235 (1-3737)

Qy   40 ValAlaAlaArgLeuAlaArgGlnHisGlnLeuTyrrValGlyValLeuGlySerLysLeuGly   59
         | | | | | | | | | | | | | | | | | | | |
Db    1 GTCGCCCGCGGACCGGCAGCACCAGCTCACGGTGGGCGTGCTCGGCAGCAAGCTGGGG        60

Qy   60 LeuGlnValValGlyLeuProAlaAspGluSerLeuProAspCysValPheValGluAsp        79
         | | | | | | | | | | | | | | | | | | |
Db   61 CTGCAGGTGGTGGAGCTGCCGGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGAC       120

Qy   80 ValAlaValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArg         99
         | | | | | | | | | | | | | | | | | |
Db  121 GTGGCCGTGTGTGCCGAGGAGACGGCGCCTCATCACCCGACCCGGGGCGCCGAGG        180

Qy  100 LysGluValAspMetMetLysGluAlaLeuGluLysGlnLeuAsnIleValGluMet       119
         | | | | | | | | | | | | | | | | | |
Db  181 AAGGAGGTTGACATGATGAAGGAAGCCATTAGAAAAACTTCAGCTAATAGTAGAGATG      240

Qy  120 LysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePhe     139
         | | | | | | | | | | | | | | | | | | |
Db  241 AAAGATGAAAATGCAACTTTAGATGCGGAGAGTGTTTATTCACAGGCAGAGAATTTTT      300

Qy  140 ValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLys     159
         | | | | | | | | | | | | | | | | | | |
Db  301 GTGGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAG   360

Qy  160 AspTyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSer     179
         | | | | | | | | | | | | | | | | | | |
Db  361 GACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGC    420

Qy  180 MetAlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLys     199
         | | | | | | | | | | | | | | | | | | |
Db  421 ATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAG    480

Qy  200 IleMetGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAla       219
         | | | | | | | | | | | | | | | | | |
Db  481 ATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCA    540

Qy  220 AlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro     239
         | | | | | | | | | | | | | | | | | | |
Db  541 GCAAACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTGCACCGACCCCG     600

Qy  240 GluGluTyrProGluSerAlaLysValTyrGluLysLysLeuLysLeuAspHisMetLeuIlePro 259
         | | | | | | | | | | | | | | | | | | | | | |
Db  601 GAAGAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACCCATATGCTGATCCCC    660

Qy  260 ValSerMetSerGluLeuGluLysValAlaAspGlyLeuLeuThrCysSerValLeuIle     279
         | | | | | | | | | | | | | | | | | | |
Db  661 GTGAGCATGTCTGAACTGGAAAAAGTGGATGGGCTGCTCACCTGCTGCTCAGTTTTAATT    720

Qy  280 AsnLysLysValAspSer                                              285
         | | | | |
Db  721 AACAAGAAAGTAGACTCC                                              738
```

```
RESULT 8
HSM808292
LOCUS       HSM808292               3908 bp    mRNA    linear   PRI 30-AUG-2003
DEFINITION  Homo sapiens mRNA; cDNA DKFZp686N2176 (from clone DKFZp686N2176).
ACCESSION   BX648145
VERSION     BX648145.1  GI:34367304
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3908)
  AUTHORS   Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  CONSRTM   The German Human cDNA Consortium
  TITLE     Direct Submission
```

JOURNAL   Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
           Neuherberg, GERMANY
COMMENT   Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
           Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
           sequenced by EMBL (European Molecular Biology Laboratories,
           Heidelberg/Germany) within the cDNA sequencing consortium of the
           German Genome Project.
           This clone (DKFZp686N2176) is available at the RZPD in Berlin.
           Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
           Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
           information about the clone and the sequencing project is available
           at http://mips.gsf.de/proj/cDNA/.
FEATURES   Location/Qualifiers
     source   1..3908
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp686N2176"
                /tissue_type="human endometrium carcinoma cell line"
                /clone_lib="686 (synonym: hlcc3). Vector pSportl_Sfi; host
                DH10B; sites SfiIA + SfiIB"
                /dev_stage="adult"
ORIGIN

Alignment Scores:
Pred. No.:   1.02e-193   Length: 3908
Score: 184.00   Matches: 184
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 64.56%   Indels: 0
DB: 9   Gaps: 0

us-09-889-733b-2 (1-285) x HSM808292 (1-3908)

```
Qy   102 ValAspMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp 121
Db   384 GTTGACATGATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGAT 443

Qy   122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGly 141
Db   444 GAAAATGCAACTTTAGATGGGGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTGGGC 503

Qy   142 LeuSerLysArgThrAsnGlnArgGlyAlaGlyIleLeuAlaAspThrPheLysAspTyr 161
Db   504 CTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTAT 563

Qy   162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
Db   564 GCAGTGTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCT 623

Qy   182 GlyProAsnLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIleMet 201
Db   624 GGGCCTAACCTGATCGCAATGGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATG 683

Qy   202 GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsn 221
Db   684 CAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAAC 743

Qy   222 CysIleTyrLeuAsnLeuProAsnLysGlyHisValLeuLeuHisArgThrProGluGlu 241
Db   744 TGTATATATCTAAATATCCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAAGAG 803

Qy   242 TyrProGluSerAlaLysValTyrGluLysLeuLeuLysAspHisMetLeuProValSer 261
Db   804 TATCCAGAAAGTGCAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGC 863

Qy   262 MetSerGluLeuGluLysValAlaAspGlyLeuLeuThrCysCysSerValLeuIleAsnLys 281
Db   864 ATGTCTGAACTGGAAAAGGTGGATGGGCTCCTCACCTGCTGCTCAGTTTTAATTAACAG 923

Qy   282 LysValAspSer 285
Db   924 AAAGTAGACTCC 935
```

```
          gene          /clone_lib="NCI_CGAP_Li9"
                        /lab_host="DH10B"
                        /note="Vector: pCMV-SPORT6"
                        1..1120
                        /gene="Ddah1"
                        /note="synonym: 2510015N06Rik"
                        /db_xref="LocusID:69219"
                        /db_xref="MGI:1916469"
          CDS           122..979
                        /codon_start=1
                        /product="dimethylarginine dimethylaminohydrolase 1"
                        /protein_id="AAH34505.1"
                        /db_xref="GI:21961624"
                        /db_xref="LocusID:69219"
                        /translation="MAGLGHPSAPGRATHAVVRAPPESLCRHALRRSQGEEVDFARAE
                        RQHELYVGULGSKLGLGQVVQLPADESLPDCVFVEDVAVVCEETALITRPGAPSRRKEV
                        DMMKEALEKLQLNIVEMKDENATLDGGDVLFTGREFFVGLSKRTNQRGAEILADTFKD
                        YAVSTVPADSLHLKSFCSMAGPNLIAIGSSESAQKALKIMQOMSDHRYDKLTVPDDM
                        AANCIYLNIPSKGHVLLHRTPEEYPESAKVYEKLKDHLLIPVSNSEMEKVDGLLTCCS
                        VFINKKIDS"
          misc_feature  155..958
                        /note="Amidinotransf; Region: Amidinotransferase. This
                        family contains glycine (EC:2.1.4.1) and inosamine
                        (EC:2.1.4.2) amidinotransferases, enzymes involved in
                        creatine and streptomycin biosynthesis respectively. This
                        family also includes arginine deiminases, EC:3.5.3.6.
                        These enzymes catalyse the reaction: arginine + H2O <=>
                        citrulline + NH3. Also found in this family is the
                        Streptococcus anti tumour glycoprotein"
                        /db_xref="CDD:pfam02274"

ORIGIN
Alignment Scores:
Pred. No.:               1.69e-107        Length:              1120
Score:                   106.00           Matches:             152
Percent Similarity:      98.70%           Conservative:        0
Best Local Similarity:   98.70%           Mismatches:          1
Query Match:             37.19%           Indels:              2
DB:                                       Gaps:                0

US-09-889-733B-2 (1-285) x BC034505 (1-1120)

Qy    65  LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaAlaValValCys    84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   314  CTGCCCGCCGACGAGAGCCTGCCCGACTGCGTCTTCGTGGAGGACGTGGCCGTCGTGTGC   373

Qy    85  GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet   104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   374  GAGGAGACGGCCCTCATCACCCGCCCCGGGGCGCCCAGCCGCAGGAGGAAGGAGGTTGACATG   433

Qy   105  MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla   124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   434  ATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATGCA   493

Qy   125  ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLys    144
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   494  ACTTTGGATGGTGGGGACGTCCTATTCACAGGACGAGAGTTTTTTGTGGGCCTTTCCAAA   553

Qy   145  ArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSer   164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   554  AGAACAAATCAACGAGGTGCTGAAATCTTTGGCTGATACTTTTAAGGACTACGCAGTCTCT   613

Qy   165  ThrValProAlaAspSerLeuHisLeuLysSerPheCysSerMetAlaGlyProAs   184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   614  ACAGTCCCGGCCGATTC-TTTGCATTTAAAGAGTTTCTGCAGCATGGCTGGCCCCAA    672

Qy   184  nLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnMe    204
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   673  CCTGATTGCAATAGGGTCCAGGAGATCTGCACAGAAGGCCCTCAAGATCATGCAACAGAT   732

Qy   204  tSerAspHisArgTyrAspLysLeuThrValProAspAsp    217
              |||||||||||||||||||||||||||||||||||||||
Db   733  GAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGAC   772
```

```
RESULT 10
LOCUS          AX683109                 3008 bp    DNA     linear   PAT 29-MAR-2003
DEFINITION     Sequence 83 from Patent EP1279744.
ACCESSION      AX683109
VERSION        AX683109.1  GI:29370146
KEYWORDS       .
SOURCE         Rattus norvegicus (Norway rat)
  ORGANISM     Rattus norvegicus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
               Rattus.
REFERENCE      1
  AUTHORS      Brookbank,R.A., Dixon,A.K., Lee,K. and Pinnock,R.D.
  TITLE        Identification and use of molecules implicated in pain
  JOURNAL      Patent: EP 1279744-A 83 29-JAN-2003;
               WARNER-LAMBERT COMPANY (US)
FEATURES             Location/Qualifiers
     source          1..3008
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
                     /note="N-G, N,G dimethylarginine dimethylaminohydrolase"
ORIGIN
Alignment Scores:
Pred. No.:               4.26e-107        Length:              3008
Score:                   106.00           Matches:             152
Percent Similarity:      98.70%           Conservative:        0
Best Local Similarity:   98.70%           Mismatches:          1
Query Match:             37.19%           Indels:              2
DB:                                       Gaps:                0

US-09-889-733B-2 (1-285) x AX683109 (1-3008)

Qy    65  LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaAlaValValCys    84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   624  CTGCCCGCCGACGAGAGCCTGCCCGACTGCGTCTTCGTGGAGGACGTGGCCGTCGTGTGC   683

Qy    85  GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet   104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   684  GAGGAGACCGCCCTCATCACCCGCCCCGGGGCGCCCAGCCGCAGGAAGGAGGTTGACATG   743

Qy   105  MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla   124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   744  ATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATGCA   803

Qy   125  ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLys    144
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   804  ACCTTAGATGGTGGGGACGTCCTATTCACAGGCAGAGAGTTTTTTGTGGGCCTTTCCAAA   863

Qy   145  ArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSer   164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   864  AGGACAAATCAACGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGACTACCCAGTTTCC   923

Qy   165  ThrValProAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs    184
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   924  ACAGTCCCGGCCGATTC-TTTGCATTTAAAGAGTTTCTGCAGCATGGCTGGCCCCAA    982

Qy   184  nLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnMe    204
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   983  CCTGATCGCAATAGGGTCCAGTGAATCTGCACAGAAGGCCCTCAAGATCATGCAACAGAT   1042

Qy   204  tSerAspHisArgTyrAspLysLeuThrValProAspAsp    217
              |||||||||||||||||||||||||||||||||||||||
Db  1043  GAGTGACCACCGTTATGACAAGCTCACTGTACCGGACGAC   1082
```

```
RESULT 11
LOCUS          D86041                   3008 bp    mRNA    linear   ROD 06-FEB-1999
DEFINITION     Rattus norvegicus mRNA for N-G,N-G-dimethylarginine
               dimethylaminohydrolase, complete cds.
```

```
ACCESSION   D86041
VERSION     D86041.1  GI:1906799
KEYWORDS    N-G,N-dimethylarginine dimethylaminohydrolase.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Kimoto,M., Sasakawa,T., Tsuji,H., Miyatake,S., Oka.T., Nio,N. and
            Ogawa.T.
  TITLE     Cloning and sequencing of cDNA encoding NG,NG-dimethylarginine
            dimethylaminohydrolase from rat kidney
  JOURNAL   Biochim. Biophys. Acta 1337 (1), 6-10 (1997)
  MEDLINE   97157050
   PUBMED   9003431
REFERENCE   2  (bases 1 to 3008)
  AUTHORS   Kimoto.M.
  TITLE     Cloning and sequencing of cDNA encoding G-N, G-N-dimethylarginine
            dimethylaminohydrolase from rat kidney
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 3008)
  AUTHORS   Kimoto.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1996) Masumi Kimoto, The University of Tokushima,
            School of Medicine, Nutrition; Kuramoto 3, Tokushima, Tokushima
            770, Japan (E-mail:kimoto@nutr.med.tokushima-u.ac.jp,
            Tel:0886-33-7088, Fax:0886-33-7089)
FEATURES             Location/Qualifiers
     source          1..3008
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /sex="female"
                     /tissue_type="kidney"
     CDS             432..1289
                     /codon_start=1
                     /note="DDAH"
                     /product="N-G,N-G-dimethylarginine dimethylaminohydrolase"
                     /protein_id="BAA18993.1"
                     /db_xref="GI:1906800"
                     /translation="MAGLSHPSVFGRATHAVVRAPPESLCRHALRRSQGEBVDFARAE
                     RQHQLYVGVLGSKLGLQVVQLPADESLPDCVFVEDVAVVCEETALITRPGAPSRRKEV
                     DMMKEALEKLQLNIVEMKDENATLDGGDVLFTGREFFVGLSKRTNQRGAEILADTFKD
                     YAVSTVPVADSLHLHKSFCSMAGPNLIAIGSSESAQKALKIMQOMSDHRYDKLTVPDDM
                     AANCIYLNIPSKGHVLLHRTPEEYPESAKYYEKLDHLLIPVSNSNEMEKVDGLLTCCS
                     VFINKKTDS"
ORIGIN
```

Alignment Scores:

Pred. No.:	4.26e-107	Length:	3008
Score:	106.00	Matches:	152
Percent Similarity:	98.70%	Conservative:	0
Best Local Similarity:	98.70%	Mismatches:	1
Query Match:	37.19%	Indels:	2
DB:		Gaps:	0

US-09-889-733B-2 (1-285) x D86041 (1-3008)

```
Qy   65 LeuProAlaAspGluSerLeuProAspCysValPheValGlyAspValAlaValValCys 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  624 CTGCCGGCGGACGAGAGCCTGCCTGACTGCGTGTTCGTCGGAGACGTGGCCGTGGTGTGC 683

Qy   85 GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet 104
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  684 GAGGAGACCGCCCTCATCACCCGCCCGGGCGCCTAGCCGCAGGAAGGAGGTTGACATG 743

Qy  105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspAspAsnAla 124
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  744 ATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATGCA 803

Qy  125 ThrLeuAspGlyGlyAspValLeuPheThrGlyYyArgGluPhePheValGlyLeuSerLys 144
```

```
Db  804 ACCTTAGATGGTGGGGACGTCCTATTCACAGGCAGAGAGTTTTTTGTGGGCCTTCCAAA 863
Qy  145 ArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSer 164
Db  864 AGGGACAAATCAACGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGACTACGCAGTTCC 923
Qy  165 ThrValProValAlaAspGly-LeuHisLeuHisLysSerPheCysSerMetAlaGlyProAs 184
Db  924 ACAGTCCCGTGGCCGATTC-TTTGCATTTAAAGAGTTTCTGCAGCATGGCTGGCCCCAA 982
Qy  184 nLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
Db  983 CCTGATCGCAATAGGGTCCAGTGAATCGCGCAGAAGGCCCTCAAGATCATGCAACAGAT 1042
Qy  204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
Db 1043 GAGTGACCACCGTTATGACAAGCTCACTGTACCGGACGAC 1082
```

```
RESULT 12
AK098405
LOCUS       AK098405                2895 bp    mRNA            linear   PRI 12-SEP-2003
DEFINITION  Homo sapiens cDNA FLJ25539 fis, clone CBR09481.
ACCESSION   AK098405
VERSION     AK098405.1  GI:21758410
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai.A., Itakura,S.,
            Fukuzumi,Y., Fujimori,Y., Morinaga,M., Suzuki,Y., Hata.H.,
            Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T.,
            Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A.,
            Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2895)
  AUTHORS   Sugano,S. and Suzuki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction and 5'-end one pass sequencing: Institute of Medical
            Science, University of Tokyo, Laboratory of Genome Structure, Human
            Genome Center; 3'-end one pass sequencing: RAB; clone selection for
            full insert sequencing: RAB and Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1..2895
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CBR09481"
                     /tissue_type="brain"
                     /clone_lib="CBR"
                     /note="cloning vector: pME18SFL3"
ORIGIN
```

Alignment Scores:

Pred. No.:		Length:	2895
Score:	1.19e-82	Matches:	97
Percent Similarity:	84.00	Conservative:	0
Best Local Similarity:	98.98%	Mismatches:	0
Query Match:	29.47%	Indels:	1
DB:	9	Gaps:	0

US-09-889-733B-2 (1-285) x AK098405 (1-2895)

```
Qy   119  MetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGlu-Ph  138
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   369  ATGAAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTTTATTCACAGGCAGAGAATTT  428

Qy   138  ePheValGlyLeuSerLysArgThrAsnGlnArgAsnGlyAlaGluIleLeuAlaAspThrPh  158
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   429  TTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTT  488

Qy   158  eLysAspTyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCy  178
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   489  TAAGGACTATGCAGTTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTG  548

Qy   178  sSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLe  198
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   549  CAGCATGGCTGGGCCTAACCTGATCGCCAATTGGGTCTCAGTAGTGAATCTGCACAGAAGGCCCT  608

Qy   198  uLysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValPro  215
               |||||||||||||||||||||||||||||||||||||||||||||||
Db   609  TAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCC  660
```

RESULT 13

```
AL162260/c
LOCUS       AL162260          104071 bp    DNA     linear   HTG 10-JUL-2001
DEFINITION  Homo sapiens chromosome 1 clone RP4-611J7 map p22.3-31.1, 8
            unordered pieces.
ACCESSION   AL162260
VERSION     AL162260.4  GI:9863624
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Mclay.K.
TITLE       Direct Submission
JOURNAL     Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 21, 2000 this sequence version replaced gi:9212914.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: dJ611J7
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator ET-amersham; 16% of reads Chemistry:
            Dye-terminator Big Dye; 83% of reads
            Consensus quality: 100691 bases at least Q40
            Consensus quality: 101774 bases at least Q30
            Consensus quality: 102472 bases at least Q20
            Insert size: 103371; sum-of-contigs
            Insert size: 112648; 4.2% error; agarose-fp
            Quality coverage: 4.76x in Q20 bases; sum-of-contigs Quality
            coverage: 4.41x in Q20 bases; agarose-fp
```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
*          1     4008:  contig of 4008 bp in length
*       4009     4108:  gap of 100 bp
*       4109    17620:  contig of 13512 bp in length
*      17621    17720:  gap of 100 bp
*      17721    29332:  contig of 11612 bp in length
```

```
*      29333    29432:  gap of 100 bp
*      29433    34595:  contig of 5163 bp in length
*      34596    34695:  gap of 100 bp
*      34696    51023:  contig of 16328 bp in length
*      51024    51123:  gap of 100 bp
*      51124    56471:  contig of 5348 bp in length
*      56472    56571:  gap of 100 bp
*      56572    62357:  contig of 5786 bp in length
*      62358    62457:  gap of 100 bp
*      62458   104071:  contig of 41614 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..104071
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p22.3-31.1"
                     /clone="RP4-611J7"
                     /clone_lib="RPCI-4"
     misc_feature    1..4008
                     /note="assembly_fragment:00515
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:left"
     misc_feature    4109..17620
                     /note="assembly_fragment:00727
                     fragment_chain:1"
     misc_feature    17721..29332
                     /note="assembly_fragment:00262
                     fragment_chain:1"
     misc_feature    29433..34595
                     /note="assembly_fragment:00257
                     fragment_chain:1"
     misc_feature    34696..51023
                     /note="assembly_fragment:00672
                     fragment_chain:1"
     misc_feature    51124..56471
                     /note="assembly_fragment:00205"
                     fragment_chain:1"
     misc_feature    56572..62357
                     /note="assembly_fragment:01056"
     misc_feature    62458..104071
                     /note="assembly_fragment:00375
                     clone_end:T7
                     vector_side:right"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                      104071
Score:                  2.66e-72    Matches:        76
Percent Similarity:     76.00       Conservative:    0
Best Local Similarity:  100.00%     Mismatches:      0
Query Match:            26.67%      Indels:          0
DB:                     100.00%     Gaps:            0
```

US-09-889-733B-2 (1-285) x AL162260 (1-104071)

```
Qy    27  GlnHisAlaLeuArgSerAlaLysGlyGluGluValAspValAlaArgAlaGluArgGln  46
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  37494  CAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTGCCCGGCGGCGGAGGGCAG  37435

Qy    47  HisGlnLeuTyrValGlyValLeuSerLysSerLeuGlyLeuGlnLysValValGluLeuPro  66
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  37434  CACCAGCTCTACGTGGGCGTGCTCAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCG  37375

Qy    67  AlaAlaAspGluSerLeuProAspCysValPheValGluAlaAspValAlaValAlaValValCysGluGlu  86
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  37374  GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTCGAGGACGTGGCCGTCGTGTGTGCGAGGAG  37315

Qy    87  ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluVal  102
               ||||||||||||||||||||||||||||||||||||||||||||||||
Db  37314  ACGGCCCTCATCACCCGACCCGGCGCCGCCAGCCCGGCGGAGAAGGAGGTA  37267
```

RESULT 14

```
LOCUS       AC019261              158405 bp    DNA     linear   HTG 12-MAR-2000
DEFINITION  Homo sapiens clone RP11-2P7, WORKING DRAFT SEQUENCE, 38 unordered
            pieces.
ACCESSION   AC019261
VERSION     AC019261.3  GI:7229998
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 158405)
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens, clone RP11-2P7
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 158405)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 12, 2000 this sequence version replaced gi:6899774.
            All repeats were identified using RepeatMasker.
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----------  Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----------  Project Information
            Center project name: L2694
            Center clone name: 2_P_7
            ----------  Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 135380 bases at least Q40
            Consensus quality: 146343 bases at least Q30
            Consensus quality: 150850 bases at least Q20
            Insert size: 173000; agarose-fp
            Insert size: 154705; sum-of-contigs
            Quality coverage: 2.8 in Q20 bases; agarose-fp
            Quality coverage: 3.2 in Q20 bases; sum-of-contigs
            --------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 38 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1607: contig of 1607 bp in length
            *     1608     1707: gap of 100 bp
            *     1708     3028: contig of 1321 bp in length
            *     3029     3128: gap of 100 bp
            *     3129     4889: contig of 1761 bp in length
            *     4890     4989: gap of 100 bp
            *     4990     6433: contig of 1444 bp in length
            *     6434     6533: gap of 100 bp
            *     6534     8085: contig of 1552 bp in length
            *     8086     8185: gap of 100 bp
            *     8186     9968: contig of 1783 bp in length
            *     9969    10068: gap of 100 bp
            *    10069    11857: contig of 1789 bp in length
            *    11858    11957: gap of 100 bp
            *    11958    14371: contig of 2414 bp in length
            *    14372    14471: gap of 100 bp
            *    14472    16047: contig of 1576 bp in length
            *    16048    16147: gap of 100 bp
            *    16148    18676: contig of 2529 bp in length
            *    18677    18776: gap of 100 bp
            *    18777    21785: contig of 3009 bp in length
            *    21786    21885: gap of 100 bp
            *    21886    24973: contig of 3088 bp in length
            *    24974    25073: gap of 100 bp
            *    25074    27441: contig of 2368 bp in length
            *    27442    27541: gap of 100 bp
            *    27542    30718: contig of 3177 bp in length
            *    30719    30818: gap of 100 bp
            *    30819    33992: contig of 3174 bp in length
            *    33993    34092: gap of 100 bp
            *    34093    36731: contig of 2639 bp in length
            *    36732    36831: gap of 100 bp
            *    36832    39283: contig of 2452 bp in length
            *    39284    39383: gap of 100 bp
            *    39384    41838: contig of 2455 bp in length
            *    41839    41938: gap of 100 bp
            *    41939    44209: contig of 2271 bp in length
            *    44210    44309: gap of 100 bp
            *    44310    47729: contig of 3420 bp in length
            *    47730    47829: gap of 100 bp
            *    47830    50393: contig of 2564 bp in length
            *    50394    50493: gap of 100 bp
            *    50494    56065: contig of 5572 bp in length
            *    56066    56165: gap of 100 bp
            *    56166    61306: contig of 5141 bp in length
            *    61307    61406: gap of 100 bp
            *    61407    65295: contig of 3889 bp in length
            *    65296    65395: gap of 100 bp
            *    65396    71547: contig of 6152 bp in length
            *    71548    71647: gap of 100 bp
            *    71648    77172: contig of 5525 bp in length
            *    77173    77272: gap of 100 bp
            *    77273    80143: contig of 2871 bp in length
            *    80144    80243: gap of 100 bp
            *    80244    84307: contig of 4064 bp in length
            *    84308    84407: gap of 100 bp
            *    84408    88430: contig of 4023 bp in length
            *    88431    88530: gap of 100 bp
            *    88531    94075: contig of 5545 bp in length
            *    94076    94175: gap of 100 bp
            *    94176   100299: contig of 6124 bp in length
            *   100300   100399: gap of 100 bp
            *   100400   107046: contig of 6647 bp in length
            *   107047   107146: gap of 100 bp
            *   107147   114514: contig of 7368 bp in length
            *   114515   114614: gap of 100 bp
            *   114615   122515: contig of 7901 bp in length
            *   122516   122615: gap of 100 bp
            *   122616   131415: contig of 8800 bp in length
            *   131416   131515: gap of 100 bp
            *   131516   138482: contig of 6967 bp in length
            *   138483   138582: gap of 100 bp
            *   138583   144615: contig of 6033 bp in length
            *   144616   144715: gap of 100 bp
            *   144716   158405: contig of 13690 bp in length.
FEATURES             Location/Qualifiers
     source          1..158405
                     /organism="Homo sapiens"
```

```
Qy    67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142713 GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAG 142772

Qy    87 ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluVal 102
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 142773 ACGGCCCTCATCACCCGACCCGGGCGCCGAGCCCGAGGAAGGAGGTA 142820

RESULT 15
AL360219/c
LOCUS       AL360219              168953 bp    DNA         linear   PRI 27-FEB-2001
DEFINITION  Human DNA sequence from clone RP11-131L23 on chromosome 1, complete
            sequence.
ACCESSION   AL360219 AC073061
VERSION     AL360219.18  GI:13169534
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 168953)
  AUTHORS   Donnelly,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On or before May 15, 2001 this sequence version replaced
            gi:9958163, gi:12718084.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e. phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-131L23 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-131L23 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true right end of clone RP11-131L23 is at 168953 in this
            sequence. The true left end of clone RP4-611J7 is at 37497 in this
            sequence. The true left end of clone RP11-290M5 is at 104198 in
            this sequence. The true right end of clone RP4-621F18 is at 100 in
            this sequence. The true right end of clone RP4-611J7 is at 140647
            in this sequence.
FEATURES             Location/Qualifiers
     source          1..168953
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-131L23"
                     /clone_lib="RPCI-11.1"
     misc_feature    13830..13855
```

```
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-2P7"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1...1607
                     /note="assembly_fragment"
     misc_feature    1708..3028
                     /note="assembly_fragment"
     misc_feature    3129...4889
                     /note="assembly_fragment"
     misc_feature    4990..6433
                     /note="assembly_fragment"
     misc_feature    6534..8085
                     /note="assembly_fragment"
     misc_feature    8186...9968
                     /note="assembly_fragment"
     misc_feature    10069...11857
                     /note="assembly_fragment"
     misc_feature    11958..14371
                     /note="assembly_fragment"
     misc_feature    14472..16047
                     /note="assembly_fragment"
     misc_feature    16148...18676
                     /note="assembly_fragment"
     misc_feature    18777...21785
                     /note="assembly_fragment"
     misc_feature    21886...24973
                     /note="assembly_fragment"
     misc_feature    25074...27441
                     /note="assembly_fragment"
     misc_feature    27542...30718
                     /note="assembly_fragment"
     misc_feature    30819...33992
                     /note="assembly_fragment"
     misc_feature    34093...36731
                     /note="assembly_fragment"
     misc_feature    36832...39283
                     /note="assembly_fragment"
     misc_feature    39384...41838
                     /note="assembly_fragment"
     misc_feature    41939...44209
                     /note="assembly_fragment"
     misc_feature    44310...47729
                     /note="assembly_fragment"
                     /vector_side:left
                     clone_end:SP6
     misc_feature    47830...50393
                     /note="assembly_fragment"
     misc_feature    50494...56065
                     /note="assembly_fragment"
     misc_feature    56166...61306
                     /note="assembly_fragment"
     misc_feature    61407...65295
                     /note="assembly_fragment"
     misc_feature    65396...71547
                     /note="assembly_fragment"
```

```
Alignment Scores:
Pred. No.:                3.95e-72
Score:                    76.00
Percent Similarity:       100.00%        Length:              158405
Best Local Similarity:    100.00%        Matches:                 76
Query Match:              26.67%         Conservative:             0
DB:                       2              Mismatches:               0
                                         Indels:                   0
                                         Gaps:                     0

US-09-889-733B-2 (1-285) x AC019261 (1-158405)

Qy    27 GlnHisAlaLeuArgSerAlaLysGlyGluGluValAspValAlaArgAlaArgGln 46
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142593 CAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTGGCGCGCCGGCGGCAG 142652

Qy    47 HisGlnLeuTyrValGlyValLeuGlySerLysSerLeuGlnLeuValValGluLeuPro 66
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142653 CACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCG 142712
```

```
misc_feature       /note="Sequence confirmed by AC073061 sequenced by WUGSC"
                   27601..27792
misc_feature       /note="Sequence confirmed by AC019261 sequenced by WIBR"
                   147415..147467
                   /note="Sequence from overlapping clone bA290M5 (AL162256).
                   Assembly confirmed by restriction digest."
misc_feature       161180..161206
                   /note="Sequence from overlapping clone bA290M5 (AL162256).
                   Assembly confirmed by restriction digest."

ORIGIN

Alignment Scores:
Pred. No.:              4.19e-72          Length:         168953
Score:                  76.00             Matches:        76
Percent Similarity:     100.00%           Conservative:   0
Best Local Similarity:  100.00%           Mismatches:     0
Query Match:            26.67%            Indels:         0
DB:                     9                 Gaps:           0

US-09-889-733B-2 (1-285) x AL360219 (1-168953)

Qy      27  GlnHisAlaLeuArgSerAlaLysGlyGluGluValAspValAlaArgValAlaAargAlaGluArgGln 46
            | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74848  CAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGACGTCGCCCGGCGGAACGGCAG 74789

Qy      47  HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74788  CACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCG 74729

Qy      67  AlaAspGluSerLeuProAspCysValPheValGluValAspAlaValValAlaValCysGluGlu 86
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74728  GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGCTGGCGGCGTGGTGTGCCGAGGAG 74669

Qy      87  ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluVal 102
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   74668  ACGGCCCTCATCACCCGACCCGGGGCGCCGAGCCCGAGGAAGGAGGTA 74621


Search completed: June  8, 2004, 02:16:41
Job time : 3367 secs
```

```
                Copyright    GenCore version 5.1.6
                Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:      June  7, 2004, 23:27:37 ; Search time 376 Seconds
                                      (without alignments)
                                      3220.044 Million cell updates/sec

Title:          US-09-889-733B-2
Perfect score:  285
Sequence:       1 MAGLGHPSAFGRATHAVVRA.........EKVDGLLTCCSVLINKKVDS 285

Scoring table:  OLIGO
                Xgapop   60.0 , Xgapext  60.0
                Xgapop   60.0 , Xgapext  60.0
                Fgapop    6.0 , Fgapext   7.0
                Delop     6.0 , Delext    7.0

Searched:       3373863 seqs, 2124099041 residues

Word size:      1

Total number of hits satisfying chosen parameters:          6745144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09889733/runat_07062004_083556_8342/app_query.fasta_1.455
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=olip2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09889733 @CGN_1_1708 @runat_07062004_083556_8342 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :     N_Geneseq_29Jan04:*
           1:  geneseqn1980s:*
           2:  geneseqn1990s:*
           3:  geneseqn2000s:*
           4:  geneseqn2001as:*
           5:  geneseqn2001bs:*
           6:  geneseqn2002s:*
           7:  geneseqn2003as:*
           8:  geneseqn2003bs:*
           9:  geneseqn2003cs:*
          10:  geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	285	100.0	858	3	AAA47655	Dimethyla
2	259	90.9	858	6	ADE71465	Human Thr
3	259	90.9	1633	4	AAH02907	Human she
4	259	90.9	1633	6	ABQ88102	Human ost
5	259	90.9	3932	7	ABZ35939	Human sec
6	259	90.9	4208	6	AAL41494	Drug meta
7	225	78.9	3778	9	ADE53981	Human pro
8	203	71.2	3651	3	AAF16077	Human pro

```
 9   69.5   198    9   858    9   ADE71463   Ade71463 Human Met
10   37.2   106   10  3008    9   ADB53014   Adb53014 Primary r
11   37.2   106   11  3008    9   ADB79843   Adb79843 Rat N-G,N
12   17.2    49 c 12 83946    6   ABQ88101   Abq88101 Human ost
13   13.3    38   13  6373    5   ABA14642   Aba14642 Human ner
14   13.3    38   14  6381    5   ABA14641   Aba14641 Human ner
15    9.5    27   15   420    3   AAC10265   Aac10265 Human sec
16    9.1    26   16   444    6   ABN89945   Abn89945 Mouse clo
17    7.0    20   17    60    6   ABN36282   Abn36282 Human spl
18    5.6    16   18   873    4   ABL03335   Abl03335 Drosophil
19    5.3    15   19   452    8   ACH27659   Ach27659 Human adu
20    5.3    15   20   472    8   ACH21905   Ach21905 Human adu
21    5.3    15   21   845    3   AAA99390   Aaa99390 Human mel
22    5.3    15   22   858    3   AAA47656   Aaa47656 Dimethyla
23    5.3    15   23  1228    3   AAD00677   Aad00677 Human Hyd
24    5.3    15   24  1376    2   AAZ52967   Aaz52967 Human pro
25    4.9    14   25   655    9   ADD34806   Add34806 Mouse mit
26    3.9    11   26   239    5   AAF68469   Aaf68469 Human lun
27    3.9    11 c 27   239    6   ABK38380   Abk38380 cDNA enco
28    3.9    11 c 28   239    6   ACA10709   Aca10709 Human lun
29    3.9    11 c 29   239    7   ABX99660   Abx99660 Lung canc
30    3.9    11   30   239   10   ADE71675   Ade71675 Human lun
31    3.9    11   31   862    6   ABQ47888   Abq47888 Oligonucl
32    3.9    11 c 32   862    6   ABQ47889   Abq47889 Oligonucl
33    3.9    11 c 33   898    6   ABQ41319   Abq41319 Oligonucl
34    3.9    11   34   898    6   ABQ41318   Abq41318 Oligonucl
35    3.5    10   35   770    6   ABS76755   Abs76755 Frog embr
36    3.5    10   36  3223    4   AAK73698   Aak73698 Human imm
37    3.5    10   37  3643    4   ABL03334   Abl03334 Drosophil
38    3.5    10   38 28772    6   ABK83555   Abk83555 Human cDN
39    3.2     9   39   456    4   AAS55803   Aas55803 Streptoco
40    3.2     9   40   641    5   AAQ05515   Aaq05515 Upstream
41    3.2     9   41   664    2   AAQ05513   Aaq05513 Upstream
42    3.2     9   42   664    2   AAQ05514   Aaq05514 Upstream
43    3.2     9   43   664    2   AAV18975   Aav18975 Nucleotid
44    3.2     9   44   774    6   ABQ86223   Abq86223 Lactobaci
45    3.2     9   45   774    9   ADE12836   Ade12836 DNA encod
```

ALIGNMENTS

```
RESULT 1
AAA47655
ID   AAA47655 standard; cDNA; 858 BP.
XX
AC   AAA47655;
XX
DT   08-NOV-2000 (first entry)
XX
DE   Dimethylarginine dimethylaminohydrolase (DDAH1) coding sequence.
XX
KW   Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH;
KW   arginine deaminase; hyperlipidemia; renal failure; hypertension;
KW   restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;
KW   ischemia reperfusion injury; septic shock; multi organ failure;
KW   arthritis; skin disorders; inflammatory cardiac disease; migraine;
KW   infection; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..858
FT                   /*tag= a
FT                   /product= "DDAH1"
XX
PN   WO200044888-A2.
XX
PD   03-AUG-2000.
XX
PF   26-JAN-2000; 2000WO-GB000226.
XX
PR   26-JAN-1999; 99GB-00001705.
PR   04-JUN-1999; 99GB-00013066.
```

```
XX
PA   (UNLO ) UNIV COLLEGE LONDON.
XX
PI   Vallance PJT, Leiper JM, Whitley GSJ, Charles IG;
XX
DR   WPI; 2000-543392/49.
DR   P-PSDB; AAB01477.
XX
PT   Novel methylarginase polypeptides and polynucleotides, used to identify
PT   modulators of them, which are used in the treatment of e.g. cancer,
PT   hypertension, and bacterial infections.
XX
PS   Claim 1; Page 55-56; 68pp; English.
XX
CC   Nucleotides encoding methylarginase polypeptides, vectors comprising
CC   these nucleotides and the polypeptides themselves can be used in
CC   medicaments for the treatment of hyperlipidemia, renal failure,
CC   hypertension, restenosis after angioplasty, atherosclerosis,
CC   complications of heart failure, schizophrenia, multiple sclerosis or
CC   cancer. Modulators of the enzyme can be used in medicaments for the
CC   treatment of ischemia-reperfusion injury of the brain or heart, cancer,
CC   lethal hypertension in severe inflammatory conditions such as septic
CC   shock or multi-organ failure, or local and systemic inflammatory
CC   disorders including arthritis, skin disorders, inflammatory cardiac
CC   disease, migraine, or microbial or bacterial infection
XX
SQ   Sequence 858 BP; 217 A; 214 C; 256 G; 171 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.71e-281      Length:         858
Score:                     285.00      Matches:        285
Percent Similarity:       100.00%      Conservative:     0
Best Local Similarity:    100.00%      Mismatches:       0
Query Match:                          Indels:           0
DB:                                   Gaps:             3

US-09-889-733B-2 (1-285) x AAA47655 (1-858)

Qy    1  MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla  20
Db    1  ATGGCCGGCCTCGGCCACCCCTCGGCCTTCGGCCGCGCCACCCACGCCGTGGTGCGGGCG  60

Qy   21  LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLeuArgSerAlaLysGlyGluValAspVal  40
Db   61  CTACCCGAGTCGCTGTGCCAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTC  120

Qy   41  AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysGlyGlyLeu  60
Db  121  GCCCGCGGAACGGCAGCAGCACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGGTGGGGCTG  180

Qy   61  GlnValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValPheValAspVal  80
Db  181  CAGGTGGTGGAGCTGCCCGGCCGACGAGAGCCCTTCCGGACTGCGTCTTCGTGGAGGACGTG  240

Qy   81  AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys  100
Db  241  GCCCTGGTGTGCGAGGAGACGACGGCCCTCATCACCCGACACCCCGGAGGCCGGAGGAAG  300

Qy  101  GluValAspMetLysGluMetLysGluLeuLysGluAsnIleValGluMetLys  120
Db  301  GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGTCTCAATATAGTAGAGATGAAA  360

Qy  121  AspGluAsnAlaThrLeuAspGlyGlyValAspValLeuPheThrGlyArgGluPheVal  140
Db  361  GATGAAAATGCAACTTTAGATGGGGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTG  420

Qy  141  GlyLeuSerArgTyrArgThrAsnGlnArgGlyAlaAlaAlaGluIleLeuAlaAspThrAspLysAsp  160
Db  421  GGCCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGAC  480

Qy  161  TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet  180
Db  481  TATGCAGTGTCCCAGTGCCCAGTGGCCAGATGGGGTTGCATTTGAAGAGTTTCTGCAGCATG  540

Qy  181  AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaAlaLeuLysIle  200
Db  541  GCTGGGCCTAACCTGATCGCAATTGGGTCTCAGTCAGTGAATCTGCACAGAGAGGCCCTTAAGATC  600

Qy  201  MetGlnGlnMetSerAspHisArgTyrAspLysTyrAspLeuThrValProAspAspIleAlaAla  220
Db  601  ATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCA  660

Qy  221  AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu  240
Db  661  AACTGTATTTATCTAAATATCCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAA  720

Qy  241  GluTyrProGluSerAlaLysValTyrGlyLysLeuLysHisMetLeuIleProVal  260
Db  721  GAGTATCCAGAAAGTGCAAAGGTTTATGGAAACTGAAGAAACCATATGCTGATCCCGTG  780

Qy  261  SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn  280
Db  781  AGCATGTCTAACTGAAAAGTGGATGGGCCTGCTCACCTGCTGCTCAGTTTTAATTAAC  840

Qy  281  LysLysValAspSer  285
Db  841  AAGAAGGTAGACTCC  855

RESULT 2
ADE71465
ID   ADE71465 standard; DNA; 858 BP.
XX
AC   ADE71465;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human Thr87 wild-type dimethylarginine dimethylaminohydrolase-1 DNA.
XX
KW   Human; dimethylarginine dimethylaminohydrolase-1; DDAH 1; chromosome 1;
KW   single nucleotide polymorphism; SNP; ADMA;
KW   asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
KW   coronary heart disease; cerebrovascular disease; hypertension; diabetes;
KW   susceptibility; genotyping; therapeutic targeting; antidiabetic;
KW   cardiant; cerebroprotective; hypotensive; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT
FT   CDS             1..858
FT                   /tag= a
FT                   /product= "Wild-type (Thr87) DDAH 1"
FT   variation       replace(260,T)
FT                   /tag= b
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
FT                   /note= "Results in Thr (C variant) or Met (T variant) at
FT                   position 87 of the protein"
XX
PN   WO2003089638-A1.
XX
PD   30-OCT-2003.
XX
PF   11-APR-2003; 2003WO-FI000274.
XX
PR   19-APR-2002; 2002US-00125456.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Valkonen V, Salonen JT, Pirskanen M, Tuomainen T, Laakso J;
PI   Laaksonen R;
XX
DR   WPI; 2003-854121/79.
DR   P-PSDB; ADE71466.
XX
PT   New nucleic acid encoding a variant dimethylarginine
PT   dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
PT   and its vascular complications, e.g. coronary or cerebrovascular disease
```

```
PT   or hypertension.
XX
PS   Disclosure; SEQ ID NO 3; 37pp; English.
XX
CC   The invention relates to a variant dimethylarginine
CC   dimethylaminohydrolase-1 (DDAH 1; ADE71464) and a nucleic acid encoding
CC   it (ADE71463). DDAHs regulate the metabolism of ADMA (asymmetrical
CC   dimethyl-arginine; a naturally occurring nitric oxide synthase inhibitor)
CC   and nitric oxide, which are relevant in disorders such as cardiovascular
CC   disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
CC   has been implicated in susceptibility to diseases such as familial
CC   combined hyperlipidaemia, premature coronary artery disease (CAD), non
CC   insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
CC   The variant DDAH 1 of the invention has Met at position 87, rather than
CC   Thr, which is caused by a C to T change in exon 1 at a position
CC   corresponding to base 260 in ADE71463 and ADE71465. Individuals hetero-
CC   or homozygous for the Thr87Met allele have an increased risk of
CC   developing cardiovascular diseases and diabetes. The invention also
CC   relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a
CC   susceptibility to a cardiovascular disease and diabetes in an individual
CC   by determining the DDAH 1 genotype; a method of treating diabetes or
CC   vascular complications of diabetes by enhancing nitric oxide
CC   availability, production or concentration; a method of targeting
CC   treatment of cardiovascular disease and diabetes in a hypertensive
CC   patient by determining the DDAH 1 genotype and treating them with a drug
CC   that affects nitric oxide availability, production or metabolism; and a
CC   transgenic animal which carries a human variant DDAH 1 nucleic acid
CC   sequence. The nucleic acid molecules and polypeptides are useful for
CC   treating cardiovascular disease (e.g., coronary heart disease;
CC   cerebrovascular disease, and hypertension), and diabetes and its vascular
CC   complications. The methods are useful for determining whether a patient
CC   will benefit from treatment with a drug which affects nitric oxide
CC   availability, production or metabolism; a drug which reduces ADMA
CC   availability or concentration; orn an agent which elevates DDAH
CC   availability or concentration (such as DDAH agonist). The methods are
CC   also useful for determining whether a patient will be at risk of adverse
CC   effects if DDAH antagonists are administered. The present sequence
CC   represents a nucleic acid encoding the "wild-type" (Thr87) DDAH 1.
XX
SQ   Sequence 858 BP; 218 A; 213 C; 258 G; 169 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score:                   1.32e-254      Length:          858
Matches:                 259.00         Matches:         259
Percent Similarity:      100.00%        Conservative:    0
Best Local Similarity:   100.00%        Mismatches:      0
Query Match:             90.88%         Indels:          0
DB:                      9              Gaps:            0

US-09-889-733B-2 (1-285) x ADE71465 (1-858)

Qy  27 GlnHisAlaLeuArgSerAlaLysGlyGluGluValAlaAspValAlaArgAlaArgGln  46
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  79 CAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTTCCGCCGCGGAACGGCAG  138

Qy  47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlnValValGlyLeuPro  66
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 139 CACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGCAGGTGGTGGGACTGCCG  198

Qy  67 AlaAspGluSerLeuProAspCysValPheValAlaValValValCysGluGlu  86
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 199 GCCGACGAGGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAG  258

Qy  87 ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAlaAspMetMetLys  106
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 259 ACGGCCCTCATCACCCGACCCGGGGCGCCGAGCCCGAGCGCCGAGGAAGGAGGTTGACATGATGAAA  318

Qy 107 GluAlaLeuGluLysGlnLeuAsnIleValGluMetLysAspGlnAsnAlaThrLeu  126
        |||||||||||||||||||||||||||||||||
Db 319 GAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTA  378

Qy 127 AspGlyGlyAspValPheThrGlyArgGluPhePheValGlyLeuSerLysSerArgThr  146
        |||||||||||||||||||||||||||||||||

Db 379 GATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTTGTGGGCTTTCCAAAAGACA  438
Qy 147 AsnGlnArgGlyAlaGlyIleLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrVal  166
        ||||||||||||||||||||||||||||||||||
Db 439 AATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTG  498
Qy 167 ProValAlaAspGlyHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle  186
        ||||||||||||||||||||||||||||||||||
Db 499 CCAGTGGCAGATGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATC  558
Qy 187 AlaIleGlySerGluSerArgAlaGlnIleLysAlaLeuLysIleMetGlnIleMetSerAsp  206
        ||||||||||||||||||||||||||||||||||
Db 559 GCAATTGGTCAGTAGATCTGCACAGAGGCCCTTAAGATCCATGCAACAGATGAGTGAC  618
Qy 207 HisArgTyrAspLysValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn  226
        ||||||||||||||||||||||||||||||||||
Db 619 CACCGGTACGACAAACTCACTGCCTGATGACTAGCAGCAAACTGTATATATCTAAAT  678
Qy 227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerArgAla  246
        ||||||||||||||||||||||||||||||||||
Db 679 ATCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCGGAAGAGTATCCAGAAAAGTGCA  738
Qy 247 LysValTyrGluLysLeuGlyLysAspHisMetLeuIleProValSerMetSerGluLeuGlu  266
        ||||||||||||||||||||||||||||||||||
Db 739 AAGGTTTATGAGAAACTGAAGGACCATATGTGCTGATCCCCGTGAGCATGTCTGAACTGGAA  798
Qy 267 LysValAspGlyLysLeuLeuThrCysSerSerValLeuIleAsnLysLysValAspSer  285
        ||||||||||||||||||||||||||||||||||
Db 799 AAGGTGGATGGGGTGCTCACCTGCTGCTCAGTTTAATTAACAAGAAAGTAGACTCC  855

RESULT 3
AAH02907
ID   AAH02907 standard; DNA; 1613 BP.
XX
AC   AAH02907;
XX
DT   15-JUN-2001 (first entry)
XX
DE   Human shear stress-response coding sequence SEQ ID NO: 67.
XX
KW   Human; shear stress-response protein; vascular disease; arteriosclerosis;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200125427-A1.
XX
PD   12-APR-2001.
XX
PF   02-OCT-2000; 2000WO-JP006840.
XX
PR   01-OCT-1999; 99JP-00280976.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
PA   (NOJI/) NOJIMA H.
XX
PI   Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
PI   Kuga T, Sekine S, Nakamura Y, Sugano S;
XX
DR   WPI; 2001-266308/27.
DR   P-PSDB; AAB90784.
XX
PT   DNA sequences, proteins encoded by them and antibodies against them
PT   useful in diagnosis and treatment of vascular disease caused by
PT   arteriosclerosis.
XX
PS   Claim 20; Page 404-407; 678pp; Japanese.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human shear stress response proteins. These are useful in the
CC   diagnosis, treatment and screening of vascular diseases caused by
CC   arteriosclerosis, including heart failure, post-PTCA restenosis and
CC   hypertension
```

XX
SQ   Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:   1633
Score: 2.45e-254   Matches: 259
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 90.88%   Indels: 0
DB: Gaps: 0

US-09-889-733B-2 (1-285) x AAH02907 (1-1633)

Qy   27 GlnHisAlaLeuArgSerAlaLysGlyGluGluValGlyGluValAspValAlaAlaArgAlaGluArgGln 46
Db   401 CAGCACCGCCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTCGCCGCCCGCGCCGAACGGCAG 460

Qy   47 HisGlnLeuTyrValGlyValLeuGlyValLeuSerLysLysGlyLeuGlnValValGluLeuPro 66
Db   461 CACCAGCTCTACGTGGGCGTGCTGGGCGTGCTGAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCG 520

Qy   67 AlaAspGluSerLeuProAspCysValPheValGluValAlaAlaValValCysGluGlu 86
Db   521 GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGCTGGTGTGCGAGGAG 580

Qy   87 ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAlaAspMetMetLys 106
Db   581 ACGGCCCTCATCACCCGACCCGGGGCGCCGAGCCGGGAGGTTGACATGATGAAA 640

Qy   107 GluAlaLeuGlyLysLeuGlnLeuAsnIleValGluMetLysSerAspGluAsnAlaThrLeu 126
Db   641 GAAGCATTAGAAAAAACTTCAGCTCAATATGTAGATAGATGAAAGATGAAAATGCAACTTTA 700

Qy   127 AspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThr 146
Db   701 GATGGCGGAGATGTTTTATTCACAGGCAGAGATTTTTGTGGGCCTTTCCAAAAGGACA 760

Qy   147 AsnGlnArgGlyAlaGluIleLeuAlaAspTyrArgLysAspTyrArgLysSerThrVal 166
Db   761 AATCAACGAGGTGCTGAAAATCTTGGCTGATACTTTTAAGGACTATCGAGTCTCCACAGTG 820

Qy   167 ProValAlaAspTyrLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle 186
Db   821 CCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATC 880

Qy   187 AlaIleGlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnMetSerAsp 206
Db   881 GCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGAC 940

Qy   207 HisArgTyrAspLysLeuThrValProAspAspProIleAlaAlaAsnCysIleTyrLeuAsn 226
Db   941 CACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAAT 1000

Qy   227 IleProAsnLysGlyHisValLeuHisArgThrProGluTyrProGluSerAla 246
Db 1001 ATCCCAAACAAAGGGCACGTCTTGCTCGCACCGGAAGAGTATCCAGAAAGTGCA 1060

Qy   247 LysValTyrGluLysLysLeuAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
Db 1061 AAGGTTTATGAGAAAACTGAAGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAA 1120

Qy   267 LysValAspGlyLeuLeuThrCysCysSerValLeuIleLeuAsnLysLysValAspSer 285
Db 1121 AAGGTGGATGGCGCTCACCTGCTGCTCAGTTTTAATTAACAAGAAGTAGACTCC 1177

RESULT 4
ABQ88102
ID   ABQ88102 standard; cDNA; 1633 BP.
XX
AC   ABQ88102;
XX
DT   18-SEP-2002 (first entry)
XX

DE   Human osteoblast differentiation related cDNA SEQ ID NO 9.
XX
KW   Human; osteoblast; stem cell differentiation; bone tissue deposition;
KW   osteoporosis; osteopathic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200250301-A2.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-US048276.
XX
PR   18-DEC-2000; 2000US-0255882P.
PR   24-APR-2001; 2001US-0285691P.
XX
PA   (GENE-) GENE LOGIC INC.
PA   (PROC) PROCTER & GAMBLE CO.
XX
PI   Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI   Mertz L;
XX
DR   WPI; 2002-557663/59.
XX
PT   Use of genes and their expression profiles associated with osteoblast
PT   differentiation for screening modulators bone formation, for diagnosing
PT   or treating e.g. osteoporosis, or as markers for the differentiation
PT   process.
XX
PS   Claim 1; SEQ ID NO 9; 78pp + Sequence Listing; English.
XX
CC   The invention relates to genes and their expression profiles are used
CC   for: (a) screening modulators of precursor stem cell differentiation into
CC   osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC   deposition of bone tissue, abnormal rate of osteoblast formation or
CC   osteoporosis; or (c) treating or monitoring treatment of the conditions
CC   cited in (b), or monitoring the progression of bone tissue deposition.
CC   Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC   osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC   induced abnormalities in bone formation or bone loss, conditions that
CC   involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC   skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC   or fibrous dysplasia. The present sequence is that of an osteoblast
CC   differentiation associated cDNA marker of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:   1633
Score: 2.45e-254   Matches: 259
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 90.88%   Indels: 0
DB: Gaps: 0

US-09-889-733B-2 (1-285) x ABQ88102 (1-1633)

Qy   27 GlnHisAlaLeuArgSerAlaLysGlyGluGluValGlyGluValAspValAlaAlaArgAlaGluArgGln 46
Db   401 CAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTCGCCGCCCGCGCCGAACGGCAG 460

Qy   47 HisGlnLeuTyrValGlyValLeuGlyValLeuSerLysLysGlyLeuGlnValValGluLeuPro 66
Db   461 CACCAGCTCTACGTGGGCGTGCTGGGCGTGCTGAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCG 520

Qy   67 AlaAspGluSerLeuProAspCysValPheValGluValAlaAlaValValCysGluGlu 86
Db   521 GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGCTGGTGTGCGAGGAG 580

Qy   87 ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAlaAspMetMetLys 106

PR   20-JUN-2001; 2001US-03000001P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI   Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI   Daughery, SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI   Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI   Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
DR   WPI; 2003-075543/07.
DR   P-PSDB; ABP75492.
XX
PT   New human secretory proteins and polynucleotides, useful for diagnosing,
PT   treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
PT   neurological disorders (e.g. Alzheimer's), or cell proliferations or
PT   cancers.
XX
PS   Claim 1; SEQ ID NO 103; 458pp + Sequence Listing; English.
XX
CC   The invention relates to a secretory polynucleotide (designated sptm)
CC   comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
CC   naturally occurring polynucleotide sequence at least 90 % identical to
CC   the polynucleotide sequence, a polynucleotide complementary to them or an
CC   RNA equivalent of them. The polypeptide or polynucleotide are useful for
CC   treating, preventing or diagnosing a disease or condition associated with
CC   the expression of functional SPTM. These are particularly useful for
CC   diagnosing, treating or preventing autoimmune/inflammatory disorders
CC   (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
CC   disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC   dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
CC   multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
CC   schizophrenia or amnesia), or cell proliferative disorders (e.g.
CC   psoriasis, polycythemia vera, or cancers including adenocarcinoma,
CC   leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
CC   breast, cervix or prostate). Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3932 BP; 1089 A; 883 C; 775 G; 1185 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   Length: 3932
Score: 5.71e-254   Matches: 259
Percent Similarity: 259.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 90.88%   Indels: 0
DB: 7   Gaps: 0

US-09-889-733B-2 (1-285) x ABZ35539 (1-3932)

Qy   27 GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln 46
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   95 CAGGACGCGCTGAGAAGCGCCAAGGGCCAAGGAGGTGGACGTGGCGCGCCGGAACGGCAG 154

Qy   47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   155 CACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGTGGGGGCTGCAGGTGGTGGAGCTGCCG 214

Qy   67 AlaAspGluSerLeuProAspCysValPheValGluValAlaValCysGluGlu 86
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   215 GCCGACGAGGACCCTTCCGGACTGCGTCCTTCGTGGAGGTGGCCGTGGTGTGCGGAGGAG 274

Qy   87 ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetLys 106
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   275 ACGGCCCTCATCACCCGACCCGGCGCGCCGAGCCCGGAGGAAGGAGGTTGACATGATGAAA 334

Qy   107 GluAlaLeuGluLysLeuGlnLeuAsnIleValGlnMetLysAspGluAsnAlaThrLeu 126
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   335 GAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTA 394

Qy   127 AspGlyGlyAspValLeuPheThrArgGlyArgGluPhePheValGlyLeuSerLysArgThr 146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   395  GATGGCGGGAGATGTTTATTCACAGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACA  454

Qy   147  AsnGlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysPheAlaValSerThrVal  166
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   455  AATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAACGACTATGCAGTCTCCACAGTG  514

Qy   167  ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle  186
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   515  CCAGTGGCAGATGGGGTTGCATTTGAAGAGTTTCTGCAGATGGCGTGGGCCTAACCTGATC  574

Qy   187  AlaIleGlySerSerGluSerArgAlaGlnIleLysAlaLeuLysIleMetGlnMetSerAsp  206
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   575  GCAATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGAC  634

Qy   207  HisArgTyrAspLysThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn  226
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   635  CACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAAT  694

Qy   227  IleProAsnLysGlyHisValLeuHisArgThrProGluGluTyrProGluSerAla  246
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   695  ATCCCCAACAAAGGGCACGTCTTGCACCGAACCCCGGAAGAGTATCCAGAAAAGTGCA  754

Qy   247  LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu  266
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   755  AAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAA  814

Qy   267  LysValAspGlyLeuLeuThrCysCysSerValLeuIleLeuAsnLysLysValAspSer  285
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   815  AAGGTGGATGGCGTGCTCACCTGCTGCTCAGTTTTAATTAACAAGAAAGTAGACTCC  871
```

RESULT 6

AAL41494

```
ID   AAL41494 standard; DNA; 4208 BP.
XX
AC   AAL41494;
XX
DT   21-NOV-2002  (first entry)
XX
DE   Drug metabolising enzyme encoding DNA - 7671089CB1.
XX
KW   Anti-HIV; antiarteriosclerotic; dermatological; cytostatic; thyromimetic;
KW   osteopathic; thrombolytic; ophthalmological; antilipaemic; hepatotropic;
KW   antidiarrhoeic; antiinflammatory; virucide; immunogenic; autoimmune;
KW   drug metabolising enzyme; DME; inflammatory; AIDS; atherosclerosis;
KW   contact dermatitis; cell-proliferative; cancer; cirrhosis; dwarfism;
KW   developmental; hypothyroidism; endocrine; osteoporosis; thrombosis;
KW   diabetes; glaucoma; keratitis; metabolic; hyperlipidaemia; diarrhoea;
KW   cystic fibrosis; gastrointestinal; gastroenteritis; liver; hepatitis;
KW   Reye's syndrome; exogenous compound; gene therapy; enzyme; human; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             200..1225
FT                   /tag= a
FT                   /product= "Drug metabolising enzyme protein"
XX
PN   WO20026654-A2.
XX
PD   29-AUG-2002.
XX
PF   14-FEB-2002; 2002WO-US004918.
XX
PR   16-FEB-2001; 2001US-0269643P.
PR   23-FEB-2001; 2001US-0271322P.
PR   16-MAR-2001; 2001US-0276767P.
PR   06-APR-2001; 2001US-0282077P.
PR   19-APR-2001; 2001US-0285447P.
PR   27-APR-2001; 2001US-0287060P.
PR   03-MAY-2001; 2001US-0288543P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
```

```
XX
PI   Astromoff A,  Au-Young J,  Baughn MR,  Ding L,  Dugan BM;
PI   Forsythe IJ,  Gietzen KJ,  Griffin JA,  Lee EA,  Lu Y,  Richardson TW;
PI   Ring HZ,  Sanjanwala MM,  Swarnakar A,  Walia NK,  Warren BA,  Xu Y;
PI   Yue H,  Zebarjadian Y;
XX
DR   WPI; 2002-674949/72.
DR   P-PSDB; AAO22798.
XX
PT   New drug metabolizing enzymes (DME) useful for diagnosing, treating and
PT   preventing diseases or conditions associated with aberrant DME
PT   expression, e.g. cancer, AIDS, atherosclerosis, diabetes, glaucoma,
PT   hepatitis, osteoporosis.
XX
PS   Claim 5; Page 164-165; 166pp; English.
XX
CC   The invention relates to an isolated polypeptide comprising 12 fully
CC   defined sequences of 81-615 amino acids given in the specification; a
CC   naturally occurring amino acid sequence at least 90% identical to, having
CC   81-599 amino acids, at least 93% identical to a sequence of 529 amino
CC   acids, or at least 97% identical to a sequence of 615 amino acids, all
CC   given in the specification; or a biologically active or immunogenic
CC   fragment of the polypeptide. The polypeptides and polynucleotides are
CC   useful in diagnosing, treating and preventing diseases or conditions
CC   associated with the decreased expression or overexpression of a drug
CC   metabolising enzyme (DME), such as autoimmune/inflammatory (e.g. AIDS,
CC   atherosclerosis, contact dermatitis) cell-proliferative (e.g. cancer,
CC   cirrhosis), developmental (e.g. dwarfism, hypothyroidism), endocrine
CC   (e.g. osteoporosis, thrombosis, diabetes), eye (e.g. glaucoma,
CC   keratitis), metabolic (e.g. hyperlipidaemia, cystic fibrosis),
CC   gastrointestinal (e.g. gastroenteritis, diarrhoea), or liver (e.g.
CC   hepatitis, Reye's syndrome) disorders. These are also useful in assessing
CC   the effects of exogenous compounds on the expression of nucleic acid and
CC   amino acid sequences of DME. The DME or its fragments are useful in
CC   screening compounds for effectiveness as agonist or antagonist of the
CC   polypeptides, or in altering the expression of the target polynucleotide
CC   and compounds that specifically bind to or modulate the activity of the
CC   polypeptide. The microarray is useful in monitoring or measuring protein-
CC   protein interactions, drug-target interactions, and gene expression
CC   profiles. The polynucleotides of the invention can be used in gene
CC   therapy. This polynucleotide sequence represents the DNA encoding a drug
CC   metabolising enzyme of the invention
XX
SQ   Sequence 4208 BP; 1092 A; 1017 C; 883 G; 1216 T; 0 U; 0 Other;
```

Alignment Scores:

```
Pred. No.:                            Length:          4208
Score:                  259.00        Matches:          259
Percent Similarity:     100.00%       Conservative:       0
Best Local Similarity:  100.00%       Mismatches:         0
Query Match:             90.88%       Indels:             0
DB:                          6        Gaps:               0
```

`6.09e-254`

US-09-889-733B-2 (1-285) x AAL41494 (1-4208)

```
Qy    27  GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln  46
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   446  CAGCACGCCCTGAGAAGCGCCAAGGGCGAGGTGGACGTGGCCCGCGCCGAACGCGCAG  505

Qy    47  HisGlnLeuTyrValGlyValLeuValSerLeuGlyLeuGlyValValGluLeuPro  66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   506  CACCAGCTCTACGTGGGCGTGCTGGTGAGCCTGGGCCTCGGGGTGGTGGAGCTGCCG  565

Qy    67  AlaAspGluSerLeuProAspCysValPheValGluAlaValAlaValValCysGluGlu  86
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   566  GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGCGGTGGCCGTGGTGTGCGAGGAG  625

Qy    87  ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLys  106
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   626  ACGGCCCTCATCACCCGACCCGGGGCGCCGAGCCCGAGGAGGAGTTGACATGATGAAA  685

Qy   107  GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysLysGluAsnAlaThrLeu  126
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
CC   containing nucleic acids by hybridising a substrate with the nucleic
CC   acids, thus forming one or more hybridisation complexes, detecting
CC   hybridisation complex formation and comparing the complexes formed with
CC   standard complexes, where differences between the standard and the sample
CC   complex formation indicate differential expression of cDNAs in the
CC   sample. The differential expression is diagnostic of prostate cancer. The
CC   invention also relates to proteins and antibodies related to the cDNAs.
CC   The combination is useful for diagnosing, treating or monitoring the
CC   progression of treatment of prostate cancer. The antibodies are useful
CC   for detecting prostate cancer. This sequence represents a human prostate
XX   cancer cDNA of the invention.
SQ   Sequence 3778 BP; 1064 A; 821 C; 719 G; 1150 T; 0 U; 24 Other;
```

Alignment Scores:
Pred. No.:
Score: 2.89e-219   Length: 3778
Percent Similarity: 225.00   Matches: 238
Best Local Similarity: 99.58%   Conservative: 0
Query Match: 99.58%   Mismatches: 0
DB: 78.95%   Indels: 1
                                             Gaps: 0

US-09-889-733B-2 (1-285) x ADE53981 (1-3778)

```
Qy   48  GlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValValGluLeuProAla  67
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2  CAGTCTCACGTGGGCGTCGTGCTGGGCAGCAAGCTGGGCCTGCAGGTGGTGGAGCTGCCGGCC  61

Qy   68  AspGluSerLeuProAspCysValPheValGluAspValAlaAlaValCysGluGluThr  87
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62  GACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGCTGGTGTGCGAGGAGACG  121

Qy   88  AlaLeuIleThrArgProGlyAlaProSerArgArgProGlnGluGluGluGluValAspMetMetLysGlu  107
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122  GCCCTCATCACCCGACCCGGGGCGCCGAGCCCGAGGAGGAGGTTGACATGATGAAAGAA  181

Qy  108  AlaLeuGluLysGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAsp  127
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182  GCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAACTTTAGAT  241

Qy  128  GlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThrAsn  147
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242  GGGGAGATGTTTATTCACAGCAGGAGAATTTTTGTGGGCCTTCCAAAAGGACAAAT  301

Qy  148  -GlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrValPr  167
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302  CCAACGAGCTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCC  361

Qy  167  oValAlaAspGlyHisLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAl  187
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  362  AGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGC  421

Qy  187  alleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHi  207
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  422  AATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCA  481

Qy  207  sArgTyrAspLysValProAspAspIleAlaIleAlaAsnCysIleTyrLeuAsnIl  227
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  482  CCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATCTAAATAT  541

Qy  227  eProAsnLysGlyHisValLeuHisArgThrProGluTyrProGluSerAlaLy  247
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  542  CCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAGAGTATCCGAAAGTGCAAA  601

Qy  247  sValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetGluGluLy  267
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  602  GGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTTGAACTGGAAA  661

Qy  267  sValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer  285
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  662  GGTGGAATGGCTGCTCACCTGCTGCCCAGTTTTTAATTAACAGAAGTAGACTCC  716

RESULT 8
```

```
Db  686  GAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAACTTTA  745
Qy  127  AspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThr  146
Db  746  GATGGCGGAGATGTTTATTCACAGCAGGAGAATTTTTGTGGGCCTTCCAAAAGGACA  805
Qy  147  AsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrVal  166
Db  806  AATCAACGAGCTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTG  865
Qy  167  ProValAlaAspGlyHisLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle  186
Db  866  CCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATC  925
Qy  187  AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp  206
Db  926  GCAATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGAC  985
Qy  207  HisArgTyrAspLysValProAspAspIleAlaIleAlaAsnCysIleTyrLeuAsn  226
Db  986  CACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATCTAAAT  1045
Qy  227  IleProAsnLysGlyHisValLeuLeuHisArgThrProGluTyrProGluSerAla  246
Db 1046  ATCCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAGAGTATCCGAAAGTGCA  1105
Qy  247  LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluGluLeuGlu  266
Db 1106  AAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAA  1165
Qy  267  LysValAspGlyLeuLeuThrCysCysSerValLeuIleLeuAsnLysLysValAspSer  285
Db 1166  AAGGTGGATGGCTGCTCACCTGCTGCCCAGTTTTTAATTAACAGAAGAAAGTAGACTCC  1222

RESULT 7
ADE53981
ID   ADE53981 standard; cDNA; 3778 BP.
XX
AC   ADE53981;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human prostate cancer cDNA #328.
XX
KW   Human; prostate cancer; ss; cDNA combination; differential expression;
KW   gene.
XX
OS   Homo sapiens.
XX
PN   US2003190640-A1.
XX
PD   09-OCT-2003.
XX
PF   29-MAY-2002; 2002US-00252157.
PR   31-MAY-2001; 2001US-0295048P.
XX
PA   (FARI/) FARIS M.
PA   (PEAR/) PEARSON C I.
XX
PI   Faris M, Pearson CI;
XX
DR   WPI; 2003-831619/77.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   prostate cancer, useful for diagnosing, treating or monitoring the
PT   progression of treatment of prostate cancer.
XX
PS   Claim 1; SEQ ID NO 328; 42pp; English.
XX
CC   The invention relates to a combination comprising a number of cDNAs
CC   expressed in prostate cancer. The invention also relates to a method for
CC   detecting differential expression of one or more cDNAs in a sample
```

AAF16077
ID AAF16077 standard; cDNA; 3651 BP.
XX
AC AAF16077;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:512.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
OS Homo sapiens.
XX
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005988.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587513/55.
DR P-PSDB; AAB56874.
XX
PT Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
XX
PS Claim 1; Page 976-977, 2338pp; English.
XX
CC AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF15506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 3651 BP; 1016 A; 801 C; 684 G; 1141 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.:   8.19e-197
Score: 203.00
Percent Similarity: 100.00%   Length: 3651
Best Local Similarity: 100.00% Matches: 203
Query Match: 71.23%   Conservative: 0
DB: 3   Mismatches: 0
                Indels: 0
                Gaps: 0

US-09-889-733B-2 (1-285) x AAF16077 (1-3651)

```
Qy    73 AspCysValPheValGluAspValAlaValAlaValCysGluThrAlaLeuIleThrArg  92
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4 GACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGACGGCCCTCATCACCCGA  63

Qy    93 ProGlyValAlaProSerArgArgLysGluValAspMetMetLysGluLeuGluLysLeu 112
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    64 CCCGGGGGCGCCGAGCGCCGGAGGAAGGAGGTTGACATGATGAAGGAGCTGGAGAAGCTT 123

Qy   113 GlnLeuAsnIleValGluMetLysAspGlyAsnAlaThrLeuAspGlyGlyAspValLeu 132
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 CAGCTCAATATAGTAGAGATGAAGGATGGAAATGCAACTTTAGATGGCGGAGATGTTTTA 183

Qy   133 PheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGlu 152
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 TTCACGGGAGAGATTTTTGTGGGCCTTCAAAAGACAAATCAACGAGGTGCTGAA 243

Qy   153 IleLeuValAspThrPhePheLysAspTyrArgAlaValSerThrValProValAlaAspGlyLeu 172
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 ATCTTGGTTGATACTTTTAAGGACTATCGAGCTGTCTCCACAGTGCCAGTGGCAGATGGGTTG 303

Qy   173 HisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleLeuAlaIleGlySerSerGlu 192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304 CATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCCGCAATTGGGTCTAGTGAA 363

Qy   193 SerAlaGlnLysAlaLeuLysIleMetGlnIleMetSerAspHisArgTyrAspLysLeu 212
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 TCTGCAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTC 423

Qy   213 ThrValProAspIleAlaAlaAsnCysIleTyrLeuAsnLysGlyHis 232
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   424 ACTGTGCCTGATGACATAGCAGCAAACTGTATATATCCCAACAAGGGCAC 483

Qy   233 ValLeuHisArgThrProGluTyrProGluSerAlaLysValTyrGluLysLeu 252
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484 GTCTTGCTCACCGAACCCCGGAAGAGTATCCGAAAGTGCAAAGGTTTATGAGAAACTG 543

Qy   253 LysAspHisMetLeuIleProValSerMetSerGlyLeuGlyLysValAspLysLeuLeu 272
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544 AAGGACCATATGCTGATCCCCGTGAGCATGTCTGGTCTGAACTGGAAAAGGTCGATGGGCTGCTC 603

Qy   273 ThrCysCys 275
          |||||||||
Db   604 ACCTGCTGT 612
```

RESULT 9
ADB71463
ID ADB71463 standard; DNA; 858 BP.
XX
AC ADB71463;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Met87 variant dimethylarginine dimethylaminohydrolase-1 DNA.
XX
KW Human; variant dimethyl-arginine dimethylaminohydrolase-1; variant DDAH 1;
KW chromosome 1; single nucleotide polymorphism; SNP; ADMA;
KW asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
KW coronary heart disease; cerebrovascular disease; hypertension; diabetes;
KW susceptibility; genotyping; therapeutic targeting; antidiabetic;
KW cardiant; cerebroprotective; hypotensive; gene therapy; gene; ds.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   1..858
FT   /tag= a
FT   /product= "Met87 variant DDAH 1"
FT variation replace(260,C)
FT   /tag= b
FT   /standard_name= "Single nucleotide polymorphism (SNP)"
FT   /note= "Results in Thr (C variant) or Met (T variant) at
FT   position 87 of the protein"
XX
PN WO2003089638-A1.
XX
PD 30-OCT-2003.
XX
PF 11-APR-2003; 2003WO-FI000274.
XX

```
PR   19-APR-2002; 2002US-00125456.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Valkonen V,  Salonen JT,  Pirskanen M,  Tuomainen T,  Laakso J;
PI   Laaksonen R;
XX
DR   WPI; 2003-854121/79.
DR   P-PSDB; ADE71464.
XX
PT   New nucleic acid encoding a variant dimethylarginine
PT   dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
PT   and its vascular complications, e.g. coronary or cerebrovascular disease
PT   or hypertension.
XX
PS   Claim 4; SEQ ID NO 1; 37pp; English.
XX
CC   The invention relates to a variant dimethylarginine
CC   dimethylaminohydrolase-1 (DDAH 1; ADE71464) and a nucleic acid encoding
CC   it (ADE71463). DDAHs regulate the metabolism of ADMA (asymmetrical
CC   dimethyl-arginine; a naturally occurring nitric oxide synthase inhibitor)
CC   and nitric oxide, which are relevant in disorders such as cardiovascular
CC   disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
CC   has been implicated in susceptibility to diseases such as familial
CC   combined hyperlipidaemia, premature coronary artery disease (CAD), non
CC   insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
CC   The variant DDAH 1 of the invention has Met at position 87, rather than
CC   Thr, which is caused by a C to T change in exon 1 at a position
CC   corresponding to base 260 in ADE71463 and ADE71465. Individuals hetero-
CC   or homozygous for the Thr87Met allele have an increased risk of
CC   developing cardiovascular diseases and diabetes. The invention also
CC   relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a
CC   susceptibility to a cardiovascular disease and diabetes in an individual
CC   by determining the DDAH 1 genotype; a method of treating diabetes or
CC   vascular complications of diabetes by enhancing nitric oxide
CC   availability, production or concentration; a method of targeting
CC   treatment of cardiovascular disease and diabetes in a hypertensive
CC   patient by determining the DDAH 1 genotype and treating them with a drug
CC   that affects nitric oxide availability, production or metabolism; and a
CC   transgenic animal which carries a human variant DDAH 1 nucleic acid
CC   sequence. The nucleic acid molecules and polypeptides are useful for
CC   treating cardiovascular disease (e.g., coronary heart disease,
CC   cerebrovascular disease, and hypertension), diabetes and its vascular
CC   complications. The methods are useful for determining whether a patient
CC   will benefit from treatment with a drug which affects nitric oxide
CC   availability, production or metabolism; a drug which reduces ADMA
CC   availability or concentration; or an agent which elevates DDAH
CC   availability or concentration (such as DDAH agonist). The methods are
CC   also useful for determining whether a patient will be at risk of adverse
CC   effects if DDAH antagonists are administered. The present sequence
CC   represents a specifically claimed nucleic acid encoding the variant
CC   (Met87) DDAH 1.
XX
SQ   Sequence 858 BP; 218 A; 212 C; 258 G; 170 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                             Length:          858
Score:                2.6e-192         Matches:         258
Percent Similarity:   198.00           Conservative:    0
Best Local Similarity: 99.23%          Mismatches:      0
Query Match:          99.23%           Indels:          1
DB:                   69.47%           Gaps:            2
                      9

US-09-889-733B-2 (1-285) x ADE71463 (1-858)

Qy   27  GlnHisAlaLeuArgSerAlaLysGlyGluGluValGluValAspValAlaArgAlaGluArgGln  46
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79  CAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGGACGTCGGCCCCGCGGAACGGCAG  138

Qy   47  HisGlnLeuTyrValGlyValLeuGlySerLysGlyLeuGlnValValGluLeuPro  66
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   139 CACCAGCTCTACGTGGGCGTGCTCGGCAGCAAGCTGGGGTCGTGGTGGAGCTGCCG  198
```

```
Qy   67  AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu  86
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199 GCCGACGAGAGCCTTCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAG  258

Qy   87  Thr-AlaLeuIleThrArgProGlyAlaProSerArgArgProGlyLysGluValAspMetMetLy  106
         |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   259 AT-GGCCCTCATCACCCGACCCGGCGCGGGCCCGAGCCCGAGGAAGGAGGTTGACATGATGAA  317

Qy   106 sGluAlaLeuGluLysGluGlnLeuAsnIleValGlumetLysAspGluAsnAlaThrIle  126
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 AGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTT  377

Qy   126 uAspGlyGlyAspValLeuPheThrGlyArgGlyPheValGlyLeuSerLysArgTh  146
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   378 AGATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGAC  437

Qy   146 rAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrVa  166
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438 AAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCAGT  497

Qy   166 lProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIl  186
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   498 GCCAGTGCAGATGGGGTTGCATTTGAAGAGTTTCTGCACCATGGCCTGGGCCTAACCTGAT  557

Qy   186 eAlaIleGlySerSerAlaGlnLysSerAlaLeuLysIleMetGlnGlnMetSerAs  206
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   558 CGCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGA  617

Qy   206 pHisArgTyrAspLysGluThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAs  226
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   618 CCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAA  677

Qy   226 nIleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAl  246
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   678 TATCCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGGAAGAGTATCCAGAAAGTGC  737

Qy   246 aLysValTyrGluLysLeuLysAspHisMetLeuLeuProValSerMetGluLeuGl  265
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   738 AAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGA  797

Qy   266 uLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer  285
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   798 AAAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTAATTAACAAGAAGTAGACTCC  855
```

```
RESULT 10
ADB53014
ID   ADB53014 standard; DNA; 3008 BP.
XX
AC   ADB53014;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3556.
XX
KW   toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW   toxicity marker; toxicity progression; drug screening;
KW   primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO200365993-A2.
XX
PD   14-AUG-2003.
XX
PF   04-FEB-2003; 2003WO-US003482.
XX
PR   04-FEB-2002; 2002US-0353171P.
PR   13-MAR-2002; 2002US-0363534P.
PR   08-APR-2002; 2002US-0370248P.
PR   10-APR-2002; 2002US-0371134P.
PR   10-APR-2002; 2002US-0371135P.
PR   10-APR-2002; 2002US-0371150P.
PR   11-APR-2002; 2002US-0371413P.
```

```
Qy   165 ThrValProValAlaAspGly-LeuHisHisLeuLysSerPheCysSerMetAlaGlyProAs  184
             ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db   924 ACAGTCCCGTGGCCGGATTC-TTTGCATTTAAAGAGTTTCTGCAGCATGGCTGGCCCCAA  982

Qy   184 nLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe  204
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   983 CCTGATCGCAATAGGGTCAGTGAATCTCGCAGAAGGCCCTCAAGATCATGCAACAGAT  1042

Qy   204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp  217
             |||||||||||||||||||||||||||||||||
Db   1043 GAGTGACCACCGTTATGACAAGCTCACTGTACCGGACGAC  1082
```

RESULT 11

ADB79843

ID   ADB79843 standard; DNA; 3008 BP.
XX
AC   ADB79843;
XX
DT   04-DEC-2003　(first entry)
XX
DE   Rat N-G,N-G-dimethylarginine dimethylaminohydrolase DNA, SEQ ID 83.
XX
KW   Analgesic; pain; streptozocin-induced diabetes; rat; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   EP1279744-A2.
XX
PD   29-JAN-2003.
XX
PF   26-JUL-2002; 2002EP-00255249.
XX
PR   27-JUL-2001; 2001GB-00018354.
PR   07-FEB-2002; 2002GB-00002910.
XX
PA   (WARN) WARNER LAMBERT CO.
XX
PI   Brooksbank RA,　Dixon AK,　Lee K,　Pinnock RD;
XX
DR   WPI; 2003-395407/38.
DR   P-PSDB; ADB79842.
XX
PT   Use of isolated gene sequences and encoded polypeptides that are
PT   upregulated in the spinal cord in response to streptozocin-induced
PT   diabetes for screening compounds for the treatment of pain, or for
PT   diagnosing pain.
XX
PS   Claim 1; Page 155; 334pp; English.
XX
CC   The present invention relates to nucleotide sequences which are useful in
CC   the screening of compounds for the treatment of pain, or for the
CC   diagnosis of pain. The nucleotide sequences are up-regulated in the
CC   spinal cord in response to streptozocin-induced diabetes. The present
CC   sequence is one such nucleotide sequence.
XX
SQ   Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:   3008
Score: 7.76e-98
Percent Similarity: 106.00   Length: 3008
Best Local Similarity: 98.70%   Matches: 152
Query Match: 98.70%   Conservative: 0
DB: 37.19%   Mismatches: 1
                          9   Indels: 2
                                           Gaps: 0

US-09-889-733B-2 (1-285) x ADB79843 (1-3008)

```
Qy    65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaValAlaValCys  84
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   624 CTGCCGGCCGACGAGAGCCTGCCTGACTGCGTGTTCGTGGAGGACGTGGCCGTGGTGTGC  683

Qy    85 GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAlaspMet  104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   684 GAGGAGACGGCCCTCATCACCCGGCCCGGGGCGCCTAGCCGCAGAAGGAGGTTGACATG  743
```

PR   19-APR-2002;　2002US-0373601P.
PR   19-APR-2002;　2002US-0373602P.
PR   22-APR-2002;　2002US-0374139P.
PR   08-MAY-2002;　2002US-0378070P.
XX
PR   09-MAY-2002;　2002US-0378652P.
PR   09-MAY-2002;　2002US-0378653P.
PR   09-MAY-2002;　2002US-0378865P.
PR   09-JUL-2002;　2002US-0394230P.
PR   09-JUL-2002;　2002US-0394253P.
PR   04-SEP-2002;　2002US-0407688P.
PR   28-JAN-2003;　2003US-0442900P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,　Porter M,　Johnson K,　Higgs B,　Castle A,　Orr M;
PI   Elashoff M;
XX
DR   WPI; 2003-731472/69.
XX
PT   Determining if a compound induces a toxic effect on a tissue or cell, for
PT   identifying hepatotoxic compounds, comprises comparing a gene expression
PT   profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT   mean values.
XX
PS   Claim 44; SEQ ID NO 3556; 874pp; English.
XX
CC   The present invention describes a method for determining whether a
CC   compound induces a toxic effect on a tissue or cell. The method comprises
CC   preparing a gene expression profile of a tissue or cell sample exposed to
CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:   3008
Score: 7.76e-98
Percent Similarity: 106.00   Length: 3008
Best Local Similarity: 98.70%   Matches: 152
Query Match: 98.70%   Conservative: 0
DB: 37.19%   Mismatches: 1
                          9   Indels: 2
                                           Gaps: 0

US-09-889-733B-2 (1-285) x ADB53014 (1-3008)

```
Qy    65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaValAlaValCys  84
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   624 CTGCCGGCCGACGAGAGCCTGCCTGACTGCGTGTTCGTGGAGGACGTGGCCGTGGTGTGC  683

Qy    85 GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAlaspMet  104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   684 GAGGAGACGGCCCTCATCACCCGGCCCGGGGCGCCTAGCCGCAGAAGGAGGTTGACATG  743

Qy   105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla  124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   744 ATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATGCA  803

Qy   125 ThrLeuAspProGlyGlyAspValAspValLeuProPheThrGlyArgGluPheValGlyLeuGlyLeuSerLys  144
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   804 ACCTTAGATGGTGGGGACGTCCTATTCACAGGCAGAGAGTTTTTGTGGGCCTTTCCAAA  863

Qy   145 ArgThrAsnGlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysAspTyrAlaValSer  164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   864 AGGACAAATCAACGAGGTGCTGAGATCTTGGCTGACATCTTCAAGGACTACGCAGTTTCC  923
```

```
Db  684 GAGGAGACCCGCCCTCATCACCGCCCGCCGGGGCGCCTAGCCGCAGGAAGGAGGTTGACATG 743
Qy  105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla 124
Db  744 ATGAAAGAGGCCTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAAATGCA 803
Qy  125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLys 144
Db  804 ACCTTAGATGGTGGGGACGTCCTATTCACAGGCAGAGAGTTTTTGTGGGCCTTTCCAAA 863
Qy  145 ArgThrArgGlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysAspTyrAlaValSer 164
Db  864 AGGACAAATCACCAGGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGACTACGCAGTTTCC 923
Qy  165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
Db  924 ACAGTCCCCGTGCCGGATTC-TTGCATTTAAAGAGTTTCTGCAGCATGGCTGGCCCCAA 982
Qy  184 nLeuIleAlaIleGlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnNMe 204
Db  983 CCTGATCGCAATAGGGTCCAGTGAATCTGCCCAGAAGGCCCTCAAGATCATGCAACAGAT 1042
Qy  204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
Db 1043 GAGTGACCACCGTTATGACAAGCTCACTGTACCGGACGAC 1082
```

RESULT 12
ABQ88101/c
ID   ABQ88101 standard; cDNA; 83946 BP.
XX
AC   ABQ88101;
XX
DT   18-SEP-2002 (first entry)
XX
DE   Human osteoblast differentiation related cDNA SEQ ID NO 8.
XX
KW   Human; osteoblast; stem cell differentiation; bone tissue deposition;
KW   osteoporosis; osteopathic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200250301-A2.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-US048276.
XX
PR   18-DEC-2000; 2000US-0255882P.
PR   24-APR-2001; 2001US-0285691P.
XX
PA   (GENE-) GENE LOGIC INC.
PA   (PROC) PROCTER & GAMBLE CO.
XX
PI   Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI   Mertz L;
XX
DR   WPI; 2002-557663/59.
XX
PT   Use of genes and their expression profiles associated with osteoblast
PT   differentiation for screening modulators bone formation, for diagnosing
PT   or treating e.g. osteoporosis, or as markers for the differentiation
PT   process.
XX
PS   Claim 1; SEQ ID NO 8; 78pp + Sequence Listing; English.
XX
CC   The invention relates to genes and their expression profiles are used
CC   for: (a) screening modulators of precursor stem cell differentiation into
CC   osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC   deposition of bone tissue, abnormal rate of osteoblast formation or
CC   osteoporosis; or (c) treating or monitoring treatment of the conditions
CC   cited in (b), or monitoring the progression of bone tissue deposition.
CC   Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC   osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC   induced abnormalities in bone formation or bone loss, conditions that
CC   involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC   skeletal disease linked to breast cancer, matocyosis, Fanconi syndrome
CC   or fibrous dysplasia. The present sequence is that of an osteoblast
CC   differentiation associated cDNA marker of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 83946 BP; 26139 A; 16620 C; 16296 G; 24891 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score: 3.0e-38   Length: 83946
Percent Similarity: 100.00%   Matches: 49
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 17.19%   Mismatches: 0
DB: Indels: 0
                                               Gaps: 0

US-09-889-733b-2 (1-285) x ABQ88101 (1-83946)

```
Qy   200 IleMetGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAla 219
Db 18610 ATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCA 18551

Qy   220 AlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro 239
Db 18550 GCAAACTGTATATATCTAAATATCCCCAACAAAGGGCACGTTCTTCTTGCTGCACGAACCCCG 18491

Qy   240 GluGluTyrProGluSerAlaLysVal 248
Db 18490 GAAGAGTATCCAGAAAGTGCAAAGTA 18464
```

RESULT 13
ABA14642
ID   ABA14642 standard; DNA; 6373 BP.
XX
AC   ABA14642;
XX
DT   23-JAN-2002 (first entry)
XX
DE   Human nervous system related polynucleotide SEQ ID NO 6973.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001334.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186150P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0209963P.
PR 26-JUL-2000; 2000US-0209964P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0230063P.
PR 14-SEP-2000; 2000US-0230064P.
PR 14-SEP-2000; 2000US-0230065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0238802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.

PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-541565/60.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
PS Disclosure; SEQ ID NO 6973; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)

```
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections. Note:The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 6373 BP; 1810 A; 1396 C; 1180 G; 1987 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:          4.42e-28          Length:            6373
Score:              38.00             Matches:           38
Percent Similarity: 100.00%           Conservative:      0
Best Local Similarity: 100.00%        Mismatches:        0
Query Match:        13.33%            Indels:            0
DB:                                   Gaps:              0

US-09-889-733B-2 (1-285) x ABA14642 (1-6373)

Qy    211 LysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLys 230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 AAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCCAACAAA 61

Qy    231 GlyHisValLeuLeuHisArgThrProGluTyrProGluSerAlaLysVal 248
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 GGGCACGTCTTGCTGCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAAGGTA 115

RESULT 14
ABA14641
ID   ABA14641 standard; DNA; 6381 BP.
XX
AC   ABA14641;
XX
DT   23-JAN-2002 (first entry)
XX
DE   Human nervous system related polynuclectide SEQ ID NO 6972.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001334.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0191123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228824P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230433P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239937P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   20-OCT-2000; 2000US-0242221P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
```

XX
SQ   Sequence 6381 BP; 1809 A; 1396 C; 1182 G; 1994 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   4.43e-28   Length: 6381
Score: 38.00   Matches: 38
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 13.33%   Indels: 0
DB: Gaps: 0

US-09-889-733B-2 (1-285) x ABA14641 (1-6381)

Qy   211 LysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLys 230
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 AAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCAAATATCCCCAACAAA 61

Qy   231 GlyHisValLeuHisArgThrProGluTyrProGluSerAlaLysVal 248
            |||||||||||||||||||||||||||||||||||||||||||||
Db   62 GGGCACGTCTTGTGCACCGAACCCGAAGAGTATCCAGAAAGTGCAAAGGTA 115

RESULT 15
AAC10265
ID   AAC10265   standard; cDNA; 420 BP.
XX
AC   AAC10265;
XX
DT   06-OCT-2000   (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 14340.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999; 99US-0122487P.
XX
PA   (GBST) GENSET.
XX
PI   Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 14340; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3', untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 420 BP; 131 A; 95 C; 94 G; 97 T; 0 U; 3 Other;

PR   08-NOV-2000; 2000US-0245523P.
PR   08-NOV-2000; 2000US-0245524P.
PR   08-NOV-2000; 2000US-0245525P.
PR   08-NOV-2000; 2000US-0245526P.
PR   08-NOV-2000; 2000US-0245527P.
PR   08-NOV-2000; 2000US-0245528P.
PR   08-NOV-2000; 2000US-0245532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250391P.
PR   01-DEC-2000; 2000US-0251160P.
PR   05-DEC-2000; 2000US-0251130P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2001US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-541565/60.
XX
PT   Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating nervous system cancers
PT   and metastases.
XX
PS   Disclosure; SEQ ID NO 6972; 1701pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC   (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infections. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences

Alignment Scores:
Pred. No.:   5.54e-18   Length: 420
Score: 27.00   Matches: 27
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 9.47%   Indels: 0
DB: 3   Gaps: 0

US-09-889-733B-2 (1-285) x AAC10265 (1-420)

Qy   102 ValAspMetMetLysGluAlaLeuGluLysGluGlnLeuAsnIleValGluMetLysAsp 121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   338 GTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGAT 397

Qy   122 GluAsnAlaThrLeuAspGly 128
             |||||||||||||||||||||
Db   398 GAAAATGCAACTTTAGATGGC 418

Search completed: June 8, 2004, 01:20:35
Job time : 411 secs

OM protein - nucleic search, using frame_plus_p2n model

Copyright GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: June 8, 2004, 01:05:17 ; Search time 2554 Seconds
(without alignments)
3332.311 Million cell updates/sec

Title: US-09-889-733B-2
Perfect score: 285
Sequence: 1 MAGLGHPSAFGRATHAVVRA.........EKVDGLLTCCSVLINKKVDS 285

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 1493109276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09889733/runat_07062004_083557_8367/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=olip2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09889733_@CGN_1_5180 @runat_07062004_083557_8367 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.		Score	% Query Match	Length	DB	ID	Description
1		171	60.0	522	9	AI751463	AI751463 cn10b04.x
2		151	53.0	889	12	BI763990	BI763990 60304987 9
3		150	52.6	455	29	AY406329	AY406329 Homo sapi
4		116	40.7	356	12	BG977444	BG977444 RC5-CIO16
5	o	111	38.9	624	10	AW514143	AW514143 hd24a02.x
6		106	37.2	535	9	AU279491	AU279491
7		106	37.2	656	13	BY742062	BY742062 BY742062
8		106	37.2	746	14	CB249709	CB249709 UI-M-EXO-
9		106	37.2	862	13	BQ899468	BQ899468 AGENCOURT
10		106	37.2	1001	13	BY709864	BY709864 BY709864
11		106	37.2	1109	11	AK010430	AK010430 Mus muscu
12		102	35.8	760	14	CA512485	CA512485 UI-R-FJ0-
13		93	32.6	702	13	BY731938	BY731938 BY731938
14	o	87	30.5	308	10	BF818963	BF818963 RC5-CI016
15		82	28.8	455	29	AY406330	AY406330 Pan trogl
16		80	28.1	685	14	CB247154	CB247154 UI-M-FI0-
17		78	27.4	777	14	CD349526	CD349526 UI-M-FY0-
18		77	27.0	385	14	CB702741	CB702741 AMGNNUC:S
19		76	26.7	556	12	BM964027	BM964027 UI-M-EQ0-
20	o	75	26.3	406	9	AI017468	AI017468 ou23c07.x
21		73	25.6	555	9	AI050531	AI050531 uc83d09.y
22		72	25.3	731	10	BB611533	BB611533 BB611533
23		70	24.6	501	29	CG617513	CG617513 OST311187
24		69	24.2	776	14	CB317151	CB317151 AGENCORT
25		68	23.9	806	9	AU051351	AU051351 AU051351
26		67	23.5	276	10	BF818967	BF818967 RC5-CI016
27		66	23.2	507	9	AI225556	AI225556 uj05e03.y
28		65	22.8	438	14	CB751364	CB751364 AMGNNUC:N
29		63	22.1	498	9	AI930353	AI930353 ul63f11.y
30		61	21.4	550	9	AA986603	AA986603 uc81h06.y
31		52	18.2	645	29	CG477727	CG477727 OST7734 M
32		52	18.2	733	14	CA750500	CA750500 UI-M-FY0-
33		51	17.9	432	10	BE864630	BE864630 UI-M-BH1-
34		50	17.5	450	14	BE787749	BE787749 AMGNNUC:N
35		50	17.5	654	9	AU296050	AU296050 AU296050
36		49	17.2	338	13	BY331431	BY331431 BY331431
37		48	16.8	639	13	BQ399610	BQ399610 NISC_mp04
38		47	16.5	486	9	AI463821	AI463821 va3la06.y
39		46	16.1	343	9	AA403666	AA403666 va3la06.r
40		46	16.1	455	29	AY406331	AY406331 Mus muscu
41	o	46	16.1	530	9	AI930086	AI930086 ul63f11.x
42	o	46	16.1	547	9	AI661738	AI661738 va3la06.x
43	o	46	16.1	548	14	CD774733	CD774733 UI-M-AQ0-
44	o	46	16.1	569	9	AI265716	AI265716 uj05e03.x
45		46	16.1	571	14	CB522710	CB522710 UI-M-GK0-

ALIGNMENTS

RESULT 1
AI751463
LOCUS   AI751463   522 bp   mRNA   linear   EST 20-JUN-2002
DEFINITION cn10b04.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
            clone NHTBC_cn10b04 random, mRNA sequence.
ACCESSION   AI751463
VERSION   AI751463.1 GI:5129727
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 522)

AUTHORS   Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
          Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
          Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
TITLE   SGAP: The Skeletal Genome Anatomy Project
JOURNAL   Unpublished (1997)
COMMENT   Contact: Libin Jia
          Medical Genetics Branch
          National Human Genome Research Institute
          10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
          Tel: 301-402-4877
          Fax: 301-496-7157
          Email: libin@helix.nih.gov
          DNA Sequencing and analyses by National Institutes of Health
          Intramural Sequencing Center (NISC).
          Plate: 10   row: b   column: 04
          Seq primer: -21M13 forward primer (ABI).
FEATURES   Location/Qualifiers
    source   1..522
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="NHTBC_cn10b04"
                  /sex="Female"
                  /tissue_type="Bone"
                  /cell_type="Trabecular Bone Cells"
                  /lab_host="SURE"
                  /clone_lib="Normal Human Trabecular Bone Cells"
                  /note="Organ: Hip; Vector: pBluescript; Site1: EcoRI;
                  Library constructed by Dr. Marian Young and Dr. Pamela
                  Gehron Robey (NIDCR)."

ORIGIN

Alignment Scores:
Pred. No.:   2.33e-156   Length: 522
Score: 171.00   Matches: 171
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 60.00%   Indels: 0
DB: 9   Gaps: 0

US-09-889-733B-2 (1-285) x AI751463 (1-522)

Qy   110 GluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGly 129
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8 GAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGAACTTTAGATGGCGGA 67

Qy   130 AspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArg 149
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 GATGTTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGACAAAATCAACGA 127

Qy   150 GlyAlaGluIleLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrValProValAla 169
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   128 GGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCA 187

Qy   170 AspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGly 189
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 GATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGG 247

Qy   190 SerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisArgTyr 209
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTAC 307

Qy   210 AspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsn 229
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GACAAACTCACTGTGCCTGATGACAGCAGCAAACTGTATATATCTAAATATCCCCAAC 367

Qy   230 LysGlyHisValLeuHisIleArgThrProGluTyrProGluSerAlaLysValTyr 249
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 AAAGGGCACGTCTTGCTGCACCGAACCCCGAAGAGTATCCAGAAAGTGCAAAGTTTAT 427

Qy   250 GluLysLeuLysAspHisMetIleLeuIleProValSerMetSerGluLeuGluLysValAsp 269
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GAGAAACTGAAGGACCCATATGCTGATCCCCGTGAGCCATGCTGGAACTGGAAAGGTGGAT 487

Qy   270 GlyLeuLeuThrCysCysSerValLeuIleAsn 280
Db   488 GGGCTGCTCACCTGCTGCTCAGTTTAATTAAC 520

RESULT 2
BI763990
LOCUS   BI763990   889 bp   mRNA   linear   EST 25-SEP-2001
DEFINITION   603040879F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5189970 5',
            mRNA sequence.
ACCESSION   BI763990
VERSION   BI763990.1   GI:15755568
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
   ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 889)
   AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
   TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
   JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11474   row: n   column: 19
            High quality sequence start: 3
            High quality sequence stop: 847.
FEATURES   Location/Qualifiers
   source   1..889
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5189970"
                  /lab_host="DH10B"
                  /clone_lib="NIH_MGC_116"
                  /note="Organ: pooled colon, kidney, stomach; Vector:
                  pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
                  source anonymous pool of 3 colons, age 26 yo male, 49 yo
                  female, 71 yo male colon; 46 yo male kidney, and pool of 2
                  stomachs, 62 yo male and 70 yo female. Library is
                  oligo-dT primed and directionally cloned (EcoRV site is
                  destroyed upon cloning). Average insert size 1.4 kb,
                  insert size range 1+3 kb. Library is normalized and
                  enriched for full-length clones and was constructed by C.
                  Gruber (Invitrogen). Research Genetics tracking code
                  023. Note: this a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.:   1.38e-136   Length: 889
Score: 151.00   Matches: 151
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 52.98%   Indels: 0
DB: 12   Gaps: 0

US-09-889-733B-2 (1-285) x BI763990 (1-889)

Qy   27 GlnHisAlaAlaLeuArgSerAlaLysGlyGluValGluValAspValAlaAlaArgAlaAlaGluArgGln 46
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   180 CAGCACGCGCTGAGAAGCCTGAGACGCCCAAGGCGCGAGGAGGTGGAAGTGGACGTGGCCGCCCGCGCCGGAACGGCAG 239

Qy   47 HisGlnLeuTyrValGlyValLeuGlySerLysArgLysGlyLeuGlnValGlyLeuGluLeuPro 66
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   240 CACCAGCTCTACGTGGGCGTGCTGGGCACAAGCTGGGCGTGCAGGTGGTGGAGCTGCCG 299

Qy   67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValAlaValCysGluGlu 86
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Pb

```
Qy  136 ArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaAlaIleLeuLeuAla 155
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3 AGAGAATTTTTGTGGGCCTTCCAAAAGGACAAATCAACGAGGTGCTGAAAATCTTGGCT 62

Qy  156 AspThrPheLysAspTyrArgAlaValSerThrValProValAlaAspGlyLeuHisLeuLys 175
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63 GATACTTTTAAGGACTATCGAGTCTCCACAGTGCCAGTGGCAGTGGGTTGCATTTGAAG 122

Qy  176 SerPheCysSerMetAlaGlyProAsnLeuIleLeuAlaGlySerGluSerAlaGln 195
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123 AGTTTCTGCAGCATGGCTGGGCCCAACCTGATCGCAATTGGTCTAGTGAATCTGCACAG 182

Qy  196 LysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLysThrValPro 215
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183 AAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCT 242

Qy  216 AspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysValHisValLeuLeu 235
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  243 GATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAAAGGGCACGTCTTGCTG 302

Qy  236 HisArgThrProGluGluTyrProGluSerArgAlaValTyrGluLysLeuLeuLysAspHis 255
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  303 CACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAAGGTTTATGAGAAAACTGAAGGACCAT 362

Qy  256 MetLeuIleProValSerMetSerGluLeuLysValAspGlyValLeuLeuThrCysCys 275
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  363 ATGCTGATCCCCGTGAGCATGTCTGAACTGAAAGTGGATGGGGTGCTCACCTGCTGC 422

Qy  276 SerValLeuIleAsnLysLysValAspSer 285
            ||||||||||||||||||||||||||||||
Db  423 TCAGTTTTAATTAACAAGAAAGTAGACTCC 452
```

```
RESULT 4
BG977444
LOCUS       BG977444                356 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION  RC5-CI0161-040101-013-D07 CI0161 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG977444
VERSION     BG977444.1  GI:14380179
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 356)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     20202663
PUBMED      10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&t2=RC5-CI0161-
            040101-013-D07&t3=2001-01-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 356.
FEATURES             Location/Qualifiers
     source          1..356
                     /organism="Homo sapiens"
```

```
Db  300 GCCGACGAGAGCCTTCCGGACGTCGTCTTCGTGGAGGACACGTGGCCGTGGTGTGCGACGAG 359
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   87 ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLys 106
Db  360 ACCGGCCCTCATCACCCGACCCCGGGCGCCGAGCCGGAGGAGGTTGACATGATGAAA 419
Qy  107 GluAlaLeuGlyLeuGlnLeuAsnIleValGlnMetLysAspGluAsnAlaThrLeu 126
Db  420 GAAGCATTAGGAAAAACTTCAGCTCAATATAGTAGATAGAGATGAAAGATGCAACTTTA 479
Qy  127 AspGlyGlyAspValLeuProPheThrGlyArgGluPheValGlyLeuSerLysArgThr 146
Db  480 GATGGCGGAGATGTTTTATTCACAGGCAGCAGAATTTTTGTGGCCCTTTCCAAAAGGACA 539
Qy  147 AsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrArgAlaValSerThrVal 166
Db  540 AATCAACGAGGTGCTGAAATCTTGGCTGACTTTAAGGACTATCGAGTCCCACAGTG 599
Qy  167 ProAlaAspGlyLeuHisLeuLysSerPhe 177
Db  600 CCAGTGGCAGATGGGTTGCATTTGAAGAGTTTT 632
```

```
RESULT 3
AY406329
LOCUS       AY406329                455 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Homo sapiens DDAH1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406329
VERSION     AY406329.1  GI:39762303
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 455)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE       Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL     Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE   2  (bases 1 to 455)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE       Direct Submission
JOURNAL     Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..455
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1..>455
                     /gene="DDAH1"
                     /locus_tag="HCM2521"
ORIGIN
```

```
Alignment Scores:
Pred. No.:
Score:                6.2e-136        Length:        455
                        150.00        Matches:       150
Percent Similarity:   100.00%        Conservative:    0
Best Local Similarity: 100.00%       Mismatches:      0
Query Match:           52.63%        Indels:          0
DB:                                   Gaps:           29

US-09-889-733B-2 (1-285) x AY406329 (1-455)
```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_xref="taxon:9606"
/clone="IMAGE:2910410"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHi19W, testis NHT, and B-cell
NCI_CGAP GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:
Pred. No.:   9.98e-98   Length: 624
Score: 111.00   Matches: 111
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 38.9%   Indels: 0
DB: 10   Gaps: 0

US-09-889-733B-2 (1-285) x AW514143 (1-624)

Qy   175 LysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAla 194
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   477 AAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCA 418

Qy   195 GlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeuThrVal 214
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   417 CAGAAGGCCCTTAAGATCATGCAACAGATGAGTGAGTGACCACCGTACGACAACTCACTGTG 358

Qy   215 ProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeu 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   357 CCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAAAGGCACGTCTTG 298

Qy   235 LeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLeuLysAsp 254
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   297 CTGCACGAACCCCGGAAGAGTATCCAGAAGTGCAAAGTTTATGAGAAACTGAAGGAC 238

Qy   255 HisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeuLeuThrCys 274
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   237 CATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGC 178

Qy   275 CysSerValLeuIleAsnLysLysValAspSer 285
         |||||||||||||||||||||||||||
Db   177 TGCTCAGTTTTAATTAACAGAAAGTAGACTCC 145

RESULT 6
AU279491
LOCUS   AU279491   535 bp   mRNA   linear   EST   31-JUL-2003
DEFINITION   AU279491 CHONS1 Homo sapiens cDNA clone CHONS1000018 5', mRNA
             sequence.
ACCESSION   AU279491
VERSION   AU279491.1   GI:28298718
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 535)
AUTHORS   Imabayashi,H., Mori,T., Gojo,S., Kiyono,T., Sugiyama,T., Irie,R.,
             Isogai,T., Hata,J., Tomoya,Y. and Umezawa,A.
TITLE   Redifferentiation of dedifferentiated chondrocytes and
             chondrogenesis of human bone marrow stromal cells via chondrosphere
             formation with expression profiling by large-scale cDNA analysis
JOURNAL   Exp. Cell Res. 288 (1), 35-50 (2003)

/mol_type="mRNA"
/db_xref="taxon:9605"
/dev_stage="Adult"
/clone_lib="CT0161"
/note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Alignment Scores:
Pred. No.:   7.08e-103   Length: 356
Score: 116.00   Matches: 116
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 40.70%   Indels: 0
DB: 12   Gaps: 0

US-09-889-733B-2 (1-285) x BG977444 (1-356)

Qy   37 GluValAspValAlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySer 56
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7 GAGGTGGACGTCGCCCGCGGCGAACGGCAGCACCAGCTCTACGTGGGCGTGCTCGGCAGC 66

Qy   57 LysLeuGlyLeuGlnValValGluLeuProAlaAspGluSerLeuProAspCysValPhe 76
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 AAGCTGGGGCTGCAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTC 126

Qy   77 ValGluAspValAlaValAlaValCysGlyGluThrAlaAlaeuIleThrArgProGlyAlaPro 96
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127 GTGGAGGACGTGCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGACCCCGGGGCGCCG 186

Qy   97 SerArgArgLysGluValAspMetMetLysGluAlaLeuGluLysGlnLeuAsnIle 116
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 187 AGCCGGAGGAAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGTCAATATA 246

Qy 117 ValGluMetLysGluAspAsnAlaThrLeuAspGlyGlyAspValPheValLeuPheThrGlyArg 136
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 247 GTAGAGATGAAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTTTATTCACAGGCAGA 306

Qy 137 GluPhePheValGlyLeuSerLysArgThrArgAsnGlnArgGlyGlyValAlaGlu 152
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 307 GAATTTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGTGCTGAA 354

RESULT 5
AW514143/c
LOCUS   AW514143   624 bp   mRNA   linear   EST   03-MAR-2000
DEFINITION   hd24a02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
             IMAGE:2910410 3' similar to TR:O94760 O94760 NG-DIMETHYLARGININE
             DIMETHYLAMINOHYDROLASE ;. mRNA sequence.
ACCESSION   AW514143
VERSION   AW514143.1   GI:7152225
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 624)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Seq primer: -40TP primer from Gibco
             High quality sequence stop: 346.
FEATURES   Location/Qualifiers
     source 1 ..624

MEDLINE   22760698
PUBMED   12878157
COMMENT   Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA Project, Sugiyama,T.; Wakamatsu,A.; Irie,R.;
            Umezawa,A.; Fukuma,M.; Kusakafi,S.; Hata,J.; Ishii,S.; Yamamoto,J.;
            Isono,Y.; Saito,K.; Nakamura,Y.; Masuho,Y.; Nagai,K.; Isogai,T.
            HRI human cDNA project; cDNA library construction & 5'-end one
            pass sequencing; Helix Research Institute.
FEATURES   Location/Qualifiers
     source   1..535
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CHONS1000018"
                     /cell_type="chondrocytes"
                     /clone_lib="CHONS1"
ORIGIN   /note="Vector: pME18SFL3"

Alignment Scores:
Pred. No.:   6.37e-93   Length: 535
Score: 106.00   Matches: 106
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 37.19%   Indels: 0
DB: 9   Gaps: 0

US-09-889-733B-2 (1-285) x AU279491 (1-535)

Qy 102 ValAspMetMetLysGluAlaAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 218 GTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGAT 277

Qy 122 GluAsnAlaThrLeuThrLeuPheThrGlyArgGluPhePheValGly 141
         ||||||||||||||||||||||||||||||||||||||||||||||
Db 278 GAAAATGCAACTTTAGATGCGGAGATGTTTTATTCACGGCAGAGAATTTTTTGTGGGC 337

Qy 142 LeuSerLysArgThrArgGlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysAspTyr 161
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 338 CTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTCGGCTGATACTTTTAAGGACTAT 397

Qy 162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 398 GCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCT 457

Qy 182 GlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMet 201
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 458 GGGCCTAACCTGATCGCGATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATG 517

Qy 202 GlnGlnMetSerAspHis 207
         ||||||||||||||||||
Db 518 CAACAGATGAGTGACCAC 535

RESULT 7
BY742062   BY742062
LOCUS   BY742062   656 bp   mRNA   linear   EST 17-DEC-2002
DEFINITION BY742062 RIKEN full-length enriched, 2 days neonate sympathetic
            ganglion Mus musculus cDNA clone 7120491N07 5', mRNA sequence.
ACCESSION   BY742062
VERSION   BY742062.1   GI:27167348
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 656)
AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,

Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp
            URL:http://genome.gsc.riken.go.jp/
            Adachi,J., Aizawa,K., Arakawa,T., Akimura,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,N., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagani,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10) 1617-1630 (2000)
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome and Genome Science Laboratory in Riken
            Genomic Sciences Center and Genome Exploration Research Group in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            RNA was provided by Dr. Akira Nakagawara (Div. of Biochemistry,
            Chiba Cancer Center Research Institute, 666-2 Nitona, Chuoh-ku,
            Chiba, 260-8717 Japan) whose assistance is gratefully acknowledged.
            Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES   Location/Qualifiers
     source   1..656
                     /organism="Mus musculus"
                     /mol_type="mRNA"

```
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="712049lN07"
/sex="mixed"
/tissue_type="sympathetic ganglion"
/dev_stage="2 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 2 days neonate
    sympathetic ganglion"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
    prepared and sequenced in Mouse Genome Encyclopedia
    Project of Genome Exploration Research Group in Riken
    Genomic Sciences Center and Genome Science Laboratory in
    RIKEN. Division of Experimental Animal Research in Riken
    contributed to prepare mouse tissues. 1st strand cDNA was
    primed with a primer [5',
    GAGAGAGAGAGAATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
    prepared by using trehalose thermo-activated reverse
    transcriptase and subsequently enriched for full-length by
    cap-trapper. Second strand cDNA was prepared with the
    primer adapter of sequence [5',
    GAGAGAGAGATTCCGAGTTAATAAATTAATCCCCCCCCCC 3']. cDNA
    was cleaved with XhoI and BamHI. Vector: a modified
    pBluescript KS(+) after bulk excision from Lambda FLC I.
    Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
    ~RNA was provided by Akira Nakagawara, Div. of
    Biochemistry, Chiba Cancer Center Research Institute,
    666-2 Nitona, Chuoh-ku, Chiba, 260-8717 Japan, whose
    assistance we gratefully acknowledge."

ORIGIN

Alignment Scores:
Pred. No.:            7.98e-93      Length:          656
Score:               106.00        Matches:         152
Percent Similarity:   98.70%       Conservative:      0
Best Local Similarity: 98.70%      Mismatches:        2
Query Match:          37.19%       Indels:            2
DB:                                Gaps:              0

US-09-889-733B-2 (1-285) x BY742062 (1-656)

Qy   65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaAlaValValCys 84
Db   12 CTGCCCCGCGACGAGAGCCTGCCCGACTGCGTGTTCGTGGAGGACGTGGCCGTCGTGTGC 71

Qy   85 GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet 104
Db   72 GAGGAGACGGCCCTCATCACCCGCCCCGGGGCGCCCAGCCCGCAGGAGGAGGTTGACATG 131

Qy  105 MetLysGluAlaLeuGluLysGluLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla 124
Db  132 ATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTGGAGATGAAAGATGAAAATGCA 191

Qy  125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLys 144
Db  192 ACTTTGGATGGTGGGGACGTCCTATTCACAGGCCAGAGAATTTTTGTGGGCCTTTCCAAA 251

Qy  145 ArgThrAsnGlnArgGlyAlaGluIleLeuLeuAlaGluThrPheLysAspTyrAlaValSer 164
Db  252 AGAACAAAATCAACGAGCGTGCTGAAATCTTGGCTGAAACTTTTAAGGACTACGCAGTCTCT 311

Qy  165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAls 184
Db  312 ACAGTCCCTGGTGGCCGATTC-TTTGCATTTAAAGAGTTTTGCAGCATGGCCGGACCCAA 370

Qy  184 nLeuIleAlaIleGlySerSerGluSerAlaGlnLysSerAlaLeuLysSerAlaLeuysIleMetGlnGlnLeuMe 204
Db  371 CCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAAGGCCCCTCAAGATCATGCAACAGAT 430

Qy  204 tSerAspHisArgTyrArgAspTyrAspLysLeuThrValProAspAsp 217
Db  431 GAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGAC 470
```

DB: 13   Gaps: 0

US-09-889-733B-2 (1-285) x BQ899468 (1-862)

```
Qy    65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspArgArgLysAlaAlaValValCys 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323 CTGCCCGCGACGAGACCCTGCCCGACTGCGTTTCGTGGAGGATCGGAGGACCTGGCCGTCGTGTGC 382

Qy    85 GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet 104
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   383 GAGGAGACGGCCCTCATCACCCGGCCCGGGGCGCCCAGCCCGCAGAAGGAGGTTGACATG 442

Qy   105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla 124
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   443 ATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATGCA 502

Qy   125 ThrLeuAspGlyGlyValAspPheThrGlyArgGluPheValGlyLeuSerLys 144
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   503 ACTTTGGATGGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAA 562

Qy   145 ArgThrAsnGlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysAspTyrAlaValSer 164
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   563 AGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTACGCAGTCTCT 622

Qy   165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
         ||||||||||||||||||    |||||||||||||||||||||||||||||||||||||||
Db   623 ACAGTCCCTGGCCGATTC-TTTGCATTTAAAGAGTTTCGCAGCATGGCCGGACCCAA 681

Qy   184 nLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682 CCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAAGGCCCTCAAGATCATGCAACAGAT 741

Qy   204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
         ||||||||||||||||||||||||||||||||||||||||
Db   742 GAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGAC 781
```

RESULT 10
BY709864
LOCUS   BY709864   1001 bp   mRNA   linear   EST 16-DEC-2002
DEFINITION   BY709864 RIKEN full-length enriched, ES cells Mus musculus cDNA
                clone 2410006N07 5', mRNA sequence.
ACCESSION   BY709864
VERSION   BY709864.1 GI:27121074
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1001)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
                Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
                Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
                Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
                Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
                Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
                Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
                Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
                Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
                Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
                Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
                Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
                Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
                Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
                Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
                Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
                Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
                Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
                Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
                Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
                Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
                Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
                Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
                Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
```

---

```
Qy   105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla 124
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   150 ATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATGCA 209

Qy   125 ThrLeuAspGlyGlyValAspPheThrGlyArgGluPheValGlyLeuSerLys 144
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   210 ACTTTGGATGGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAA 269

Qy   145 ArgThrAsnGlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysAspTyrAlaValSer 164
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   270 AGAACAAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTACGCAGTCTCT 329

Qy   165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
         ||||||||||||||||||    |||||||||||||||||||||||||||||||||||||||
Db   330 ACAGTCCCTGGCCGATTC-TTTGCATTTAAAGAGTTTCGCAGCATGGCCGGACCCAA 388

Qy   184 nLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   389 CCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAAGGCCCTCAAGATCATGCAACAGAT 448

Qy   204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
         |||||||||||||||||||||||
Db   449 GAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGAC 488
```

RESULT 9
BQ899468
LOCUS   BQ899468   862 bp   mRNA   linear   EST 16-AUG-2002
DEFINITION   AGENCOURT_8750357 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6332815
                5', mRNA sequence.
ACCESSION   BQ899468
VERSION   BQ899468.1 GI:22291482
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 862)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-r@mail.nih.gov
                Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
                Ph.D.
                cDNA Library Preparation: ResGen, Invitrogen Corp
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLAM13791 row: a column: 08
                High quality sequence stop: 704.
FEATURES   Location/Qualifiers
     source   1..862
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6332815"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_130"
                     /note="Organ: oocysts; Vector: pCMV-SPORT6.1, Site 1:
                     EcoRV; Site_2: NotI; Cloned unidirectionally. Primer:
                     Oligo dT. Average insert size 1.95 kb. Constructed by
                     ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

Alignment Scores:
Pred. No.:   1.08e-92   Length: 862
Score: 106.00   Matches: 152
Percent Similarity: 98.70%   Conservative: 0
Best Local Similarity: 98.70%   Mismatches: 1
Query Match: 37.19%   Indels: 2
```

Best Local Similarity: 98.70%　　　　Mismatches:　1
Query Match: 37.19%　　　　Indels:　2
DB: 13　　　　Gaps:　0

US-09-889-733B-2 (1-285) x BY709864 (1-1001)

```
Qy    65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCys  84
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 CTGCCCGCCGACGAGAGCCTGCCCGACTGCGTGTTCGTGGAGGACGTGGCCGTCGTGTGC 387

Qy    85 GluGluThrArgAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet 104
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388 GAGGAGACGGCGGCCCTCATCACCCGCCCCGGGGCCCCAGCCGCAGGAGGAGGTTGACATG 447

Qy   105 MetLysGluAlaLeuGlyLeuGlnLeuAsnIleValGluMetLysGluAsnAla 124
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   448 ATGAAGGAGGCTTTGGGAAAACTTCAGTCAACATAGTAGAGATGAAAGATGAAAAATGCA 507

Qy   125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLys 144
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508 ACTTTGGATGGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTTGTGGGCCTTTCCAAA 567

Qy   145 ArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrProLysAspTyrAlaValSer 164
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568 AGAACAAATCAACGAGGTGCTGAAACTTGGCTGATACTTTAAGGACTACGCAGTCTCT 627

Qy   165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
         |||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   628 ACAGTCCCTGTGGCCGATTC-TTTGCATTTAAAGAGTTTCTGCAGCATGGCCGGACCCAA 686

Qy   184 nLeuIleAlaIleGlySerSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   687 CCTGATTGCAATAGGGTCCAGCCAATCTGCACAGAAGGCGCTCCAAGATCATCAACAGAT 746

Qy   204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
         ||||||||||||||||||||||||||||||||||||||||
Db   747 GAGTGACCATCGTTATGACAAGCTCACTGTACCGACGAC 786
```

RESULT 11
AK010430
AK010430
LOCUS　　　　AK010430　　　　1109 bp　　　　mRNA　　　　linear　　　　HTC 20-SEP-2003
DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library,
　　　　clone:2410006N07 product:dimethylarginine dimethylaminohydrolase 1,
　　　　full insert sequence.
ACCESSION　AK010430
VERSION　　AK010430.1　GI:12845867
KEYWORDS　 HTC; CAP trapper.
SOURCE　　 Mus musculus (house mouse)
　ORGANISM Mus musculus
　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
　AUTHORS Carninci,P. and Hayashizaki,Y.
　TITLE　 High-efficiency full-length cDNA cloning
　JOURNAL Meth. Enzymol. 303, 19-44 (1999)
　MEDLINE 99279253
　PUBMED　 10349636
REFERENCE 2
　AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
　　　　Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
　TITLE　 Normalization and subtraction of cap-trapper-selected cDNAs to
　　　　prepare full-length cDNA libraries for rapid discovery of new genes
　JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
　MEDLINE 20499374
　PUBMED　 11042159
REFERENCE 3
　AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
　　　　Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
　　　　Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
　　　　Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
　　　　Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
　　　　Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
　　　　Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

　　TITLE　 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
　　　　Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
　　　　Rogers,J., Birney,E. and Hayashizaki,Y.
　　　　Analysis of the mouse transcriptome based on functional annotation
　　　　of 60,770 full-length cDNAs
　JOURNAL Nature 420, 563-573 (2002)
　MEDLINE 22354683
　PUBMED　 12466851
COMMENT　 Contact: Yoshihide Hayashizaki
　　　　Laboratory for Genome Exploration Research Group, RIKEN Genomic
　　　　Sciences Center (GSC), Yokohama Institute
　　　　The Institute of Physical and Chemical Research (RIKEN)
　　　　1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
　　　　Tel: 81-45-503-9222
　　　　Fax: 81-45-503-9216
　　　　Email: genome-res@gsc.riken.go.jp,
　　　　URL:http://genome.gsc.riken.go.jp/
　　　　Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
　　　　Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
　　　　Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
　　　　Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
　　　　Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
　　　　Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
　　　　Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
　　　　Direct Submission
　　　　Computational Analysis of Full-Length Mouse cDNAs Compared with
　　　　Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
　　　　Normalization and subtraction of cap-trapper-selected cDNAs to
　　　　prepare full-length cDNA libraries for rapid discovery of new
　　　　genes. Genome Res. 10 (10), 1617-1630 (2000)
　　　　RIKEN integrated sequence analysis (RISA) system--384-format
　　　　sequencing pipeline with 384 multicapillary sequencer. Genome Res.
　　　　10 (11), 1757-1771 (2000)
　　　　Computer-based methods for the mouse full-length cDNA
　　　　encyclopedia: real-time sequence clustering for construction of a
　　　　nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
　　　　cDNA library was prepared and sequenced in Mouse Genome
　　　　Encyclopedia Project of Genome Exploration Research Group in Riken
　　　　Genomic Sciences Center and Genome Science Laboratory in RIKEN.
　　　　Division of Experimental Animal Research in Riken contributed to
　　　　prepare mouse tissues.
　　　　Please visit our web site (http://genome.gsc.riken.go.jp) for
　　　　further details.
FEATURES　　　　Location/Qualifiers
　　source　　1..1001
　　　　　/organism="Mus musculus"
　　　　　/mol_type="mRNA"
　　　　　/strain="C57BL/6J"
　　　　　/db_xref="taxon:10090"
　　　　　/clone="2410006N07"
　　　　　/cell_type="ES cells"
　　　　　/lab_host="SOLR"
　　　　　/clone_lib="RIKEN full-length enriched, ES cells"
　　　　　/note="Site 1: XhoI; Site 2: BamHI; cDNA library was
　　　　　prepared and sequenced in Mouse Genome Encyclopedia
　　　　　Project of Genome Exploration Research Group in Riken
　　　　　Genomic Sciences Center and Genome Science Laboratory in
　　　　　RIKEN. Division of Experimental Animal Research in Riken
　　　　　contributed to prepare mouse tissues. 1st strand cDNA was
　　　　　primed with a primer [5'
　　　　　GAGAGAGAGAGAGAAGAGATCCAAGAGCTCTTTTTTTTTTTTTTTTTTVN 3'], cDNA was
　　　　　prepared by using trehalose thermo-activated reverse
　　　　　transcriptase and subsequently enriched for full-length by
　　　　　cap-trapper. cDNA went through one round of normalization
　　　　　to Rot = 5.0 and subtraction to Rot = 25.0. Second strand
　　　　　cDNA was prepared with the primer adapter of sequence [5'
　　　　　GAGAGAGAGATTCTCGAGTTAATTAAATTAAATCCCCCCCCCCCCCCC 3'].

ORIGIN
Alignment Scores:
Pred. No.:　1.28e-92　　Length:　1001
Score:　106.00　　Matches:　152
Percent Similarity:　98.70%　　Conservative:　0

```
TITLE       RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
JOURNAL     Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE     20530913
PUBMED      11076861
REFERENCE   4
AUTHORS     The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
TITLE       Functional annotation of a full-length mouse cDNA collection
JOURNAL     Nature 409, 685-690 (2001)
REFERENCE   5
AUTHORS     The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
TITLE       Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL     Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1109)
AUTHORS     Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE       Direct Submission
JOURNAL     Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAGGATCCAAGAGCTCTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGAGATTCTCGAGTTAATAAATAATCCCCCCCCCCCCCCC 3'], cDNA was cleaved
            with XhoI and SstI. Cloning sites; 5' end: XhoI, 3' end: SstI.
            Host: SOLR.
FEATURES             Location/Qualifiers
     source          1..1109
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2410006N07"
                     /db_xref="MGI:MGI:1901307"
                     /db_xref="taxon:10090"
                     /clone="2410006N07"
                     /cell_type="ES cells"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
     CDS             137..994
                     /note="unnamed protein product; dimethylarginine
                     dimethylaminohydrolase 1 (MGD|MGI:1916469, GB|AK010430,
                     evidence: BLASTN, 100%, match=1109)
                     putative"
                     /codon_start=1
                     /protein_id="BAB26932.1"
                     /db_xref="GI:12845868"
                     /translation="MAGLGHPSAFGRATHAVVRAPPESLCRHALRRSQGEVDPARAE
                     RQHELYVGVLGSKLGLQVVCLPADESLPDCVFVEDVAVVCEETALITRPGAPSRRKEV
                     DMKKFALEKLQLNIVEMKDENATLDGGDVLFTGREFFVGLSKRTNQRGAEILADTFKD
```

```
                                                                YAVSTVPVADSLHLKSFCSMAGPNLIAIGSSESAQKALKIMQQMSDHRYDKLTVPDDM
                                                                AANCIYLNIPSKGHVLLHRTPEEYPESAKVYEKLKDHLLIPVSNSEMEKVDGLLTCCS
                                                                VFINKKIDS"

ORIGIN

Alignment Scores:
Pred. No.:             1                                         Length:          1109
Score:                 1.43e-92                                  Matches:          152
Percent Similarity:    106.00                                   Conservative:       0
Best Local Similarity: 98.70%                                   Mismatches:         1
Query Match:           98.70%                                    Indels:             2
DB:                    37.19%                                    Gaps:               0

US-09-889-733b-2 (1-285) x AK010430 (1-1109)

Qy   65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaAlaValValCys  84
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  329 CTGCCCGCGACGAGGAGCCTGCCCGACTGCGTGTTCGTGGAGGACGTGGCCGTCGTGTGC 388

Qy   85 GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet 104
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  389 GAGGAGACGGCCCTCATCACCCGCCCGGGCGCCCCAGCCGCAGGAGGAGGTTGACATG 448

Qy  105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla 124
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  449 ATGAAGAAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATGCA 508

Qy  125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheePheValGlyLeuSerLys 144
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  509 ACTTTGGATGGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTGTGGGGCCTTTCAAA 568

Qy  145 ArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspThrArgAlaValSer 164
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  569 AGAACAAATCAACGAGGTGCTGAAATCTGGCTGATACTTTTAAGGACTACGCAGTCTCT 628

Qy  165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
        |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  629 ACAGTCCCTGTGGCCGATTC-TTTGCATTTAAAGAGTTTCTGCAGCATGGCCGGACCCCA 687

Qy  184 nLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  688 CCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAGGCCCTCAAGATCATGCAACAGAT 747

Qy  204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
        |||||||||||||||||||||||||||||||||||||
Db  748 GAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGAC 787
```

```
RESULT 12
CA512485
LOCUS       CA512485                 760 bp    mRNA    linear    EST 15-NOV-2002
DEFINITION  UI-R-FJ0-cpz-1-06-0-UI.r1 UI-R-FJ0 Rattus norvegicus cDNA clone
            UI-R-FJ0-cpz-1-06-0-UI 5', mRNA sequence.
ACCESSION   CA512485
VERSION     CA512485.1  GI:25003439
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 760)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road, 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
```

```
Qy      208 gTyrAspLysLeuThrValProAspAsp 217
            ||||||||||||||||||||||
Db      421 TTATGACAAGCTCACTGTACCGACGAC 448
```

RESULT 13
BY731938
LOCUS   BY731938   702 bp   mRNA   linear   EST 17-DEC-2002
DEFINITION BY731938 RIKEN full-length enriched, 8 cells embryo Mus musculus
            cDNA clone E860038F15 5', mRNA sequence.
ACCESSION   BY731938
VERSION   BY731938.1 GI:27145065
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 702)
AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Pletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pilla,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp
            URL:http://genome.gsc.riken.go.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            RIKEN integrated sequence analysis (RISA) system--384-format
```

```
FEATURES     Location/Qualifiers
source       1..760
             /organism="Rattus norvegicus"
             /mol_type="mRNA"
             /strain="Sprague-Dawley"
             /db_xref="taxon:10116"
             /clone="UI-R-FJ0-cpz-l-06-0-UI"
             /tissue_type="embryo"
             /dev_stage="embryo"
             /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
             /clone_lib="UI-R-FJ0"
             /note="Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I;
             UI-R-FJ0 is a cDNA library containing the following
             tissue(s): rat embryo. The library was constructed
             according to Bonaldo, Lennon and Soares, Genome Research,
             6:791-806, 1996. First strand cDNA synthesis was primed
             with an oligo-dT primer containing a Not I site. Double
             stranded cDNA was ligated to an EcoR I adaptor, digested
             with Not I, and cloned directionally into pT7T3-Pac
             vector. The oligonucleotide used to prime the synthesis of
             first-strand cDNA contains a library tag sequence that is
             located between the Not I site and the (dt)18 tail. The
             sequence tag for this library is CATCTCTACT. This library
             was created for the University of Iowa Program for Rat
             Gene Discovery and Mapping (Val Sheffield, Bento Soares
             and Tom Casavant)"

ORIGIN

Alignment Scores:
Pred. No.:                             Length:          760
Score:                  7.52e-89       Matches:         148
Percent Similarity:     102.00         Conservative:      0
Best Local Similarity:  98.67%         Mismatches:        1
Query Match:            98.67%         Indels:            2
DB:                     35.79%         Gaps:              0
                        14

US-09-889-733B-2 (1-285) x CA512485 (1-760)
```

```
Qy   69  GluSerLeuProAspCysValPheValGluAspValAlaValValCysGluThrAla 88
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2  GAGAGCCGCTGCCTGACTGCGTGTTCGTGGACGACGTGGCCGTGGTGTGCCAGGAGACCGCC 61

Qy   89  LeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLysGluAla 108
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62  CTCATCACCCGCCCCGGGGCCCTAGCCCGCAGGAGGAAGGAGGTTGACATGATGAAAGAGGCT 121

Qy  109  LeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGly 128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122  TTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATGCAACCTTAGATGGT 181

Qy  129  GlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThrAsnGln 148
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182  GGGGACGTCCATTCACAGGCAGAGAGTTTTTTGTGGGCCTTTCCAAAAGGACAAAATCAA 241

Qy  149  ArgGlyAlaGluIleLeuAlaAspThrProLysAspTyrAlaValSerThrValProVal 168
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242  CGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGACTACGGAGTTTCCACAGTCCCCGTG 301

Qy  169  AlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIle 188
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302  GCCGATTC-TTTGCATTTAAAGAGTTTCTGCAGCATGGCTGGCCCCAACCTGATCGCAAT 360

Qy  188  eGlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisAr 208
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  AGGGTCCAGTGAATCTGCCGCAGAAGGCCCTCAAGATCATGCAACAGAGTGAGTGACCACCG 420
```

JOURNAL   sequence tags
             Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE   20202663
PUBMED   10737800
COMMENT   Contact: Simpson A.J.G.
             Laboratory of Cancer Genetics
             Ludwig Institute for Cancer Research
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel: +55-11-2704922
             Fax: +55-11-2707001
             Email: asimpson@ludwig.org.br
             This sequence was derived from the FAPESP/LICR Human Cancer Genome
             Project. This entry can be seen in the following URL
             (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&t2=RC5-CIO160-
             221200-012-A08&t3=2000-12-22&t4=1)
             Seq primer: puc 18 forward
             High quality sequence start: 11
             High quality sequence stop: 308.
FEATURES   Location/Qualifiers
     source 1..308
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /dev_stage="Adult"
             /clone_lib="CIO160"
             /note="Organ: colon ins; Vector: puc18; Site 1: SmaI;
             Site 2: SmaI; A mini-library was made by cloning products
             derived from ORESTES PCR (U.S. Letters Patent application
             No. 196,716 - Ludwig Institute for Cancer Research)
             profiles into the pUC 18 vector. Reverse transcription of
             tissue mRNA and cDNA amplification were performed under
             low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:   Length: 308
Score: 1.19e-74   Matches: 87
Percent Similarity: 87.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 30.53%   Gaps: 0
                      10

US-09-889-733B-2 (1-285) x BF818963 (1-308)

Qy   61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheGluAspVal   80
Db   308 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTG 249

Qy   81 AlaValValCysGluGlyThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
Db   248 GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGGCCCGGGGCGCCGAGCCGGAGGAAG 189

Qy   101 GluValAspMetMetLysGluAlaLeuGluValGlnLeuAsnIleValGluMetLys 120
Db   188 GAGGTTGACATGATGAAAGAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAA 129

Qy   121 AspGluAsnAlaThrLeuAspGlyArgGlyAspValLeuPheThrGlyArgGluPheVal 140
Db   128 GATGAAAATGCAACTTTAGATGGCGGAGATGTTTTATTCACAGGCAGCAGAGAATTTTTGTG 69

Qy   141 GlyLeuSerLysArgThrAsn 147
Db   68 GGCCTTTCCAAAAGGACAAAT 48

RESULT 15
AY406330
LOCUS   AY406330   455 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION   Pan troglodytes DDAH1 gene, VIRTUAL TRANSCRIPT, partial sequence,
             genomic survey sequence.
ACCESSION   AY406330
VERSION   AY406330.1 GI:39762304
KEYWORDS   GSS.

---

sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
FEATURES   Location/Qualifiers
     source 1..702
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="C57BL/6J"
             /db_xref="taxon:10090"
             /clone="E860038F15"
             /cell_type="8 cells"
             /dev_stage="8 cells embryo"
             /clone_lib="RIKEN full-length enriched, 8 cells embryo"
ORIGIN

Alignment Scores:
Pred. No.:   Length: 702
Score: 4.17e-80   Matches: 93
Percent Similarity: 93.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 32.63%   Gaps: 0
                      13

US-09-889-733B-2 (1-285) x BY731938 (1-702)

Qy   65 LeuProAlaAspGluSerLeuProAspCysValPheGluAspValAlaValValValCys   84
Db   375 CTGCCCGCCGACGAGAGCCTGCCCGACTGCGTGTTCGTGGAGGACGTGGCCGTCGTGTGC 434

Qy   85 GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet 104
Db   435 GAGGAGACGGCCCTCATCACCCGGCCCGGGGCGCCCAGCCGCAGGAAGGAGGTTGACATG 494

Qy   105 MetLysGluAlaLeuGluValGlnLeuAsnIleValGluMetLysAspGluAsnAla 124
Db   495 ATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATGCA 554

Qy   125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuGlyLeuSerLys 144
Db   555 ACTTTGGATGGTGGGGACGTCCTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAA 614

Qy   145 ArgThrAsnGlnArgGlyAlaAlaGlyIleLeuLeuAlaAspThr 157
Db   615 AGACAAAATCAACGAGGTGCTGAAATCTTGGCTGATACT 653

RESULT 14
BF818963/c
LOCUS   BF818963   308 bp   mRNA   linear   EST 13-JAN-2001
DEFINITION   RC5-CIO160-221200-012-A08 CIO160 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF818963
VERSION   BF818963.1 GI:12156591
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 308)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
             Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
             Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
             Brunstein,A., deOliveira,A.P.S., Bucher,P., Jongeneel,C.V.,
             O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
             Simpson,A.J.
  TITLE   Shotgun sequencing of the human transcriptome with ORF expressed

```
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM    Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 455)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE       Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL     Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE   2  (bases 1 to 455)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE       Direct Submission
JOURNAL     Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..455
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1..>455
                     /gene="DDAH1"
                     /locus_tag="HCM2521"
ORIGIN

Alignment Scores:
Pred. No.:                 1.39e-69            Length:        455
Score:                     82.00               Matches:       149
Percent Similarity:        98.68%              Conservative:  0
Best Local Similarity:     98.68%              Mismatches:    0
Query Match:               28.77%              Indels:        1
DB:                        29                  Gaps:          2
                                                              0

US-09-889-733B-2 (1-285) x AY406330 (1-455)

Qy  136 ArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAla 155
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3 AGAGAATTTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCT  62

Qy  156 AspThrPheLysAspTyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLys 175
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63 GATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAG 122

Qy  176 SerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGln 195
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123 AGTTTCTGCAGCATGGCCGGGCCTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAG 182

Qy  196 LysAlaLeuLysIleMetSerAspHisArgTyrAspLysLeuThrValPro 215
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183 AAGGCCCTTAAGATCATGAGTGACCACCGCTATGACAAACTCACTGTGCCT 242

Qy  216 AspAspIle-AlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLe 235
            ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db  243 GATGACGT-AGCAGCAAACTGTATATATCCCACAAAGGGCACGTCTTGCT 301

Qy  235 uHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHi 255
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302 GCACGGACCCCCGGAAGAGTATCCAGAAAGTGCAAAGTTTATGAGAAACTGAAGGACCCA 361

Qy  255 sMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCy 275
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  362 TATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACCTGCTG 421

Qy  275 sSerValLeuIleAsnLysLysValAspSer 285
            ||||||||||||||||||||||||||||||
```

```
Search completed: June  8, 2004, 02:57:28
Job time : 2570 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 01:20:43 ; Search time 94 Seconds
                                    (without alignments)
                                    1682.564 Million cell updates/sec

Title: US-09-889-733B-2
Perfect score: 285
Sequence: 1 MAGLGHPSAFGRATHAVVRA.........EKVDGLLTCCSVLINKKVDS 285

Scoring table: OLIGO
        Xgapop 60.0 , Xgapext 60.0
        Ygapop 60.0 , Ygapext 60.0
        Fgapop   6.0 , Fgapext   7.0
        Delop   6.0 , Delext   7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360442

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE1=frame+_p2n.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool_p/US09889733/runat_07062004_083558_8387/app_query.fasta_1.455
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=olip2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09889733 @CGN_1_1105 @runat_07062004_083558_8387 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :   Issued_Patents_NA:*
      1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
      2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
      3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
      4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
      5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
      6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	15	5.3	1397	4	US-09-023-655-37	Sequence 37, Appl
2	11	3.9	239	4	US-09-702-705-404	Sequence 404, App
3	11	3.9	239	4	US-09-736-457-404	Sequence 404, App
4	11	3.9	239	4	US-09-614-124B-404	Sequence 404, App
5	11	3.9	239	4	US-09-671-325-404	Sequence 404, App
6	11	3.9	239	4	US-09-589-184-404	Sequence 404, App
7	9	3.2	664	4	US-08-455-755-1	Sequence 1, Appli
8	9	3.2	664	2	US-07-910-260-1	Sequence 1, Appli
9	9	3.2	664	5	PCT-US91-00878-1	Sequence 1, Appli
10	9	3.2	774	4	US-09-724-623-43	Sequence 43, Appl
11	9	3.2	1089	4	US-09-134-000C-1376	Sequence 1376, Ap
12	9	3.2	1212	4	US-09-107-532A-112	Sequence 112, App

13	9	3.2	1596	4	US-09-134-001C-737	Sequence 737, App
14	9	3.2	4253	4	US-08-956-171E-118	Sequence 118, App
15	9	3.2	7360	1	US-08-286-740-1	Sequence 1, Appli
16	9	3.2	7360	5	PCT-US95-09576-1	Sequence 1, Appli
17	9	3.2	10011	5	US-08-961-527-76	Sequence 76, Appl
18	8	2.8	363	6	5182210-19	Patent No. 518221O
19	8	2.8	477	6	US-09-489-039A-4729	Sequence 4729, Ap
20	8	2.8	509	6	5182210-15	Patent No. 518221O
21	8	2.8	516	6	US-09-252-991A-14927	Sequence 14927, A
22	8	2.8	537	6	US-09-252-991A-2859	Sequence 2859, Ap
23	8	2.8	539	6	5182210-13	Patent No. 518221O
24	8	2.8	591	6	5182210-17	Patent No. 518221O
25	8	2.8	729	6	5182210-21	Patent No. 518221O
26	8	2.8	1335	4	US-09-252-991A-14446	Sequence 14446, A
27	8	2.8	1386	4	US-09-489-039A-4802	Sequence 4802, Ap
28	8	2.8	1476	4	US-09-252-991A-2659	Sequence 2659, Ap
29	8	2.8	1488	4	US-09-252-991A-2759	Sequence 2759, Ap
30	8	2.8	1522	4	US-09-443-067-25	Sequence 25, Appl
31	8	2.8	1593	4	US-09-252-991A-5394	Sequence 5394, Ap
32	8	2.8	2181	4	US-09-443-067-19	Sequence 19, Appl
33	8	2.8	2191	4	US-09-127-219B-1	Sequence 1, Appli
34	8	2.8	2298	4	US-09-252-991A-3207	Sequence 3207, Ap
35	8	2.8	5190	4	US-09-134-000C-2270	Sequence 2270, Ap
36	8	2.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
37	8	2.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
38	7	2.5	48	1	US-07-885-689A-8	Sequence 8, Appli
39	7	2.5	210	4	US-09-540-236-898	Sequence 898, App
40	7	2.5	216	3	US-09-795-430-2	Sequence 2, Appli
41	7	2.5	216	4	US-09-355-700-2	Sequence 2, Appli
42	7	2.5	216	4	US-08-601-132-38	Sequence 38, Appl
43	7	2.5	216	4	US-08-671-573B-38	Sequence 38, Appl
44	7	2.5	241	4	US-09-857-063-24	Sequence 24, Appl
45	7	2.5	244	4	US-09-857-063-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-023-655-37
; Sequence 37, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/023,655
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:

```
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1397 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: HMC1NOT01
CLONE: 003437
US-09-023-655-37

Alignment Scores:
Pred. No.:                                Length:        1397
Score:                   8.49e-07         Matches:       15
Percent Similarity:      15.00            Conservative:  0
Best Local Similarity:   100.00%          Mismatches:    0
Query Match:             100.00%          Indels:        0
DB:                      5.26%            Gaps:          0
                         4

US-09-889-733B-2 (1-285) x US-09-023-655-37 (1-1397)

Qy   130 AspValLeuPheThrGlyArgGluPheValGlyLeuSerLys 144
             |||||||||||||||||||||||||||||||||||||||
Db   716 GACGTTCTCTTCACCGGCCGGGAGTTTTTCGTAGGCCTCTCCAAA 760


RESULT 2
US-09-702-705-404/c
; Sequence 404, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
APPLICANT:  Wang, Tongtong
APPLICANT:  Bangur, Chaitanya S.
APPLICANT:  Lodes, Michael A.
APPLICANT:  Fanger, Gary
APPLICANT:  Vedvick, Tom
APPLICANT:  Carter, Darrick
APPLICANT:  Retter, Marc
APPLICANT:  Mannion, Jane
APPLICANT:  Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 404
LENGTH: 239
TYPE: DNA
ORGANISM: Homo sapien
US-09-702-705-404

Alignment Scores:
Pred. No.:                                Length:        239
Score:                   0.00359          Matches:       11
Percent Similarity:      11.00            Conservative:  0
Best Local Similarity:   100.00%          Mismatches:    0
Query Match:             100.00%          Indels:        0
DB:                      3.86%            Gaps:          0
                         4

US-09-889-733B-2 (1-285) x US-09-702-705-404 (1-239)

Qy   130 AspValLeuPheThrGlyArgGluPheVal 140
             ||||||||||||||||||||||||||||
Db    35 GACGTTCTCTTCACCGGCCGGGAGTTTTTCGTA 3


RESULT 3
US-09-736-457-404/c
; Sequence 404, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
APPLICANT:  Wang, Tongtong
APPLICANT:  Bangur, Chaitanya S.
APPLICANT:  Lodes, Michael A.
APPLICANT:  Fanger, Gary
APPLICANT:  Vedvick, Tom
APPLICANT:  Carter, Darrick
APPLICANT:  Retter, Marc
APPLICANT:  Mannion, Jane
APPLICANT:  Fan, Liqun
APPLICANT:  Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 404
LENGTH: 239
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-404

Alignment Scores:
Pred. No.:                                Length:        239
Score:                   0.00359          Matches:       11
Percent Similarity:      11.00            Conservative:  0
Best Local Similarity:   100.00%          Mismatches:    0
Query Match:             100.00%          Indels:        0
DB:                      3.86%            Gaps:          0
                         4

US-09-889-733B-2 (1-285) x US-09-736-457-404 (1-239)

Qy   130 AspValLeuPheThrGlyArgGluPheVal 140
             ||||||||||||||||||||||||||||
Db    35 GACGTTCTCTTCACCGGCCGGGAGTTTTTCGTA 3


RESULT 4
US-09-614-124B-404/c
; Sequence 404, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
APPLICANT:  Wang, Tongtong
APPLICANT:  Bangur, Chaitanya S.
APPLICANT:  Lodes, Michael A.
APPLICANT:  Fanger, Gary
APPLICANT:  Vedvick, Tom
APPLICANT:  Carter, Darrick
APPLICANT:  Retter, Marc
APPLICANT:  Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C9
CURRENT APPLICATION NUMBER: US/09/614,124B
CURRENT FILING DATE: 2001-07-11
NUMBER OF SEQ ID NOS: 1668
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 404
LENGTH: 239
TYPE: DNA
ORGANISM: Homo sapien
US-09-614-124B-404

Alignment Scores:
Pred. No.:                                Length:        239
Score:                   0.00359          Matches:       11
Percent Similarity:      11.00            Conservative:  0
Best Local Similarity:   100.00%          Mismatches:    0
Query Match:             100.00%          Indels:        0
DB:                      3.86%            Gaps:          0
                         4

US-09-889-733B-2 (1-285) x US-09-614-124B-404 (1-239)
```

```
Qy  130 AspValLeuPheThrGlyArgGluPhePheVal 140
            |||||||||||||||||||||||||
Db   35 GACGTTCTCTTCACCGGCCGCGGGAGTTTTTCGTA 3
```

RESULT 5
US-09-671-325-404/c
; Sequence 404, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-404

Alignment Scores:			
Pred. No.:	0.00359	Length:	239
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.86%	Indels:	0
DB:	4	Gaps:	0

US-09-889-733B-2 (1-285) x US-09-671-325-404 (1-239)

```
Qy  130 AspValLeuPheThrGlyArgGluPhePheVal 140
            |||||||||||||||||||||||||
Db   35 GACGTTCTCTTCACCGGCCGCGGGAGTTTTTCGTA 3
```

RESULT 6
US-09-589-184-404/c
; Sequence 404, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-589-184-404

Alignment Scores:			
Pred. No.:	0.00359	Length:	239
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.86%	Indels:	0
DB:	4	Gaps:	0

US-09-889-733B-2 (1-285) x US-09-589-184-404 (1-239)

```
Qy  130 AspValLeuPheThrGlyArgGluPhePheVal 140
            |||||||||||||||||||||||||
Db   35 GACGTTCTCTTCACCGGCCGCGGGAGTTTTTCGTA 3
```

RESULT 7
US-08-455-755-1
; Sequence 1, Application US/0455755
; Patent No. 5736345
; GENERAL INFORMATION:
; APPLICANT: Mather, Jennie P.
; APPLICANT: Roberts, Penelope E.
; TITLE OF INVENTION: LUNG CELL LINE AND METHODS OF USE
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,755
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/910260
; FILING DATE: 16-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/00878
; FILING DATE: 08-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/479130
; FILING DATE: 09-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 585P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 664 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-455-755-1

Alignment Scores:			
Pred. No.:	1.66	Length:	664
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.16%	Indels:	0
DB:	1	Gaps:	0

US-09-889-733B-2 (1-285) x US-08-455-755-1 (1-664)

```
Qy   54 LeuGlySerLysLeuGlyLeuGlnVal 62
```

```
Db      609 CTCGGTTCTAAGCTTGGGCTGCAGGTC 635

RESULT 8
US-07-910-260-1
; Sequence 1, Application US/07910260
; Patent No.5830685
; GENERAL INFORMATION:
;   APPLICANT: Mather, Jennie P.
;   APPLICANT: Roberts, Penelope E.
;   TITLE OF INVENTION: Lung Cell Line and Methods of Use
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Genentech, Inc.
;     STREET: 460 Point San Bruno Blvd
;     CITY: South San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94080
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: patin (Genentech)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/07/910,260
;     FILING DATE: 16-JUL-1992
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US91/00878
;     FILING DATE: 08-FEB-1991
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 07/479130
;     FILING DATE: 9-FEB-1990
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Hasak, Janet E.
;     REGISTRATION NUMBER: 28,616
;     REFERENCE/DOCKET NUMBER: 585P1
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415/225-1896
;     TELEFAX: 415/952-9881
;     TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 664 bases
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
US-07-910-260-1

Alignment Scores:
Pred. No.:              1.66        Length:        664
Score:                  9.00        Matches:          9
Percent Similarity:   100.00%       Conservative:     0
Best Local Similarity:100.00%       Mismatches:       0
Query Match:            3.16%       Indels:           0
DB:                        2        Gaps:             0

US-09-889-733B-2 (1-285) x US-07-910-260-1 (1-664)

Qy    54 LeuGlySerLysLeuGlyLeuGlnVal 62
         ||||||||||||||||||||||
Db   609 CTCGGTTCTAAGCTTGGGCTGCAGGTC 635
US-07-910-260-1

RESULT 9
PCT-US91-00878-1
; Sequence 1, Application PC/TUS9100878
; GENERAL INFORMATION:
;   APPLICANT: Genentech, Inc.
;   APPLICANT: Mather, Jennie P.
;   APPLICANT: Roberts, Penelope E.
;   TITLE OF INVENTION: Lung Cell Line And Methods of Use
;   NUMBER OF SEQUENCES: 1
```

```
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Genentech, Inc.
;     STREET: 460 Point San Bruno Boulevard
;     CITY: South San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94080
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: patin (Genentech)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US91/00878
;     FILING DATE: 19910208
;     CLASSIFICATION: 435/240
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: U.S.N. 07/479,130
;     FILING DATE: 9-FEB-1990
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Benson, Robert H.
;     REGISTRATION NUMBER: 30,446
;     REFERENCE/DOCKET NUMBER: 585P1
;   TELECOMUNICATION INFORMATION:
;     TELEPHONE: 415/266-1489
;     TELEFAX: 415/952-9881
;     TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 664 bases
;     TYPE: NUCLEIC ACID
;     STRANDEDNESS: single
;     TOPOLOGY: linear
PCT-US91-00878-1

Alignment Scores:
Pred. No.:              1.66        Length:        664
Score:                  9.00        Matches:          9
Percent Similarity:   100.00%       Conservative:     0
Best Local Similarity:100.00%       Mismatches:       0
Query Match:            3.16%       Indels:           0
DB:                        5        Gaps:             0

US-09-889-733B-2 (1-285) x PCT-US91-00878-1 (1-664)

Qy    54 LeuGlySerLysLeuGlyLeuGlnVal 62
         ||||||||||||||||||||||
Db   609 CTCGGTTCTAAGCTTGGGCTGCAGGTC 635

RESULT 10
US-09-724-623-43
; Sequence 43, Application US/09724623
; Patent No. 6476209
; GENERAL INFORMATION:
;   APPLICANT: Glenn, Matthew
;   APPLICANT: Lubbers, Mark W
;   APPLICANT: Dekker, James
;   TITLE OF INVENTION: Polynucleotides, materials incorporating
;     them, and methods for using them.
;   FILE REFERENCE: 1048U1
;   CURRENT APPLICATION NUMBER: US/09/724,623
;   CURRENT FILING DATE: 2000-11-28
;   NUMBER OF SEQ ID NOS: 124
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 43
;     LENGTH: 774
;     TYPE: DNA
;     ORGANISM: Lactobacillus rhamnosus
US-09-724-623-43

Alignment Scores:
Pred. No.:              1.94        Length:        774
Score:                  9.00        Matches:          9
```

```
Percent Similarity:      100.00%        Conservative:    0
Best Local Similarity:   100.00%        Mismatches:      0
Query Match:             3.16%          Indels:          0
DB:                      4              Gaps:            0

US-09-889-733B-2 (1-285) x US-09-724-623-43 (1-774)

Qy    107 GluAlaLeuGluLysLeuGlnLeuAsn 115
          |||||||||||||||||||||||||||
Db    574 GAGGCATTGGAAAAGTTGCAACTCAAC 600

RESULT 11
US-09-134-000C-1376
; Sequence 1376, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1376
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1376

Alignment Scores:
Pred. No.:               2.76           Length:          1089
Score:                   9.00           Matches:         9
Percent Similarity:      100.00%        Conservative:    0
Best Local Similarity:   100.00%        Mismatches:      0
Query Match:             3.16%          Indels:          0
DB:                      4              Gaps:            0

US-09-889-733B-2 (1-285) x US-09-134-000C-1376 (1-1089)

Qy    107 GluAlaLeuGluLysLeuGlnLeuAsn 115
          |||||||||||||||||||||||||||
Db    988 GAGGCATTAGAAAAATTACAAATTAAAT 1014

RESULT 12
US-09-107-532A-112
; Sequence 112, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm  and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                     ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;   STREET: 100 Beaver Street
;   CITY: Waltham
;   STATE: Massachusetts
;   COUNTRY: USA
;   ZIP: 02354
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: CD/ROM ISO9660
;   COMPUTER: PC
;   OPERATING SYSTEM: <Unknown>
;   SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/107,532A
;   FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/085,598
;   FILING DATE: 14 May 1998
```

```
;   APPLICATION NUMBER: 60/051571
;   FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
;   NAME: Arinello, Pamela Deneke
;   REGISTRATION NUMBER: 40,489
;   REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (781)893-5007
;   TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 112:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1212 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;     ORGANISM: Enterococcus faecium
;   FEATURE:
;     NAME/KEY:   misc_feature
;     LOCATION: (B) LOCATION 1...1212
;     SEQUENCE DESCRIPTION: SEQ ID NO: 112:
US-09-107-532A-112

Alignment Scores:
Pred. No.:               3.09           Length:          1212
Score:                   9.00           Matches:         9
Percent Similarity:      100.00%        Conservative:    0
Best Local Similarity:   100.00%        Mismatches:      0
Query Match:             3.16%          Indels:          0
DB:                      4              Gaps:            0

US-09-889-733B-2 (1-285) x US-09-107-532A-112 (1-1212)

Qy    107 GluAlaLeuGluLysLeuGlnLeuAsn 115
          |||||||||||||||||||||||||||
Db    310 GAAGCTTTAGAAAAACTGCAATTGAAT 336

RESULT 13
US-09-134-001C-737
; Sequence 737, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 737
; LENGTH: 1596
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-737

Alignment Scores:
Pred. No.:               4.1            Length:          1596
Score:                   9.00           Matches:         9
Percent Similarity:      100.00%        Conservative:    0
Best Local Similarity:   100.00%        Mismatches:      0
Query Match:             3.16%          Indels:          0
DB:                      4              Gaps:            0

US-09-889-733B-2 (1-285) x US-09-134-001C-737 (1-1596)

Qy    107 GluAlaLeuGluLysLeuGlnLeuAsn 115
```

```
Db    1063  GAAGCTTTAGAAAAATTACAACTTAAT 1089
             ||||||   ||||||||||||||||||||||
```

RESULT 14
US-08-956-171E-118

; Sequence 118, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
     APPLICANT: Charles Kunsch
              Gil H. Choi
              Patrick S. Dillon
              Craig A. Rosen
              Steven C. Barash
              Michael R. Fannon
     TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
     NUMBER OF SEQUENCES: 5256
     CORRESPONDENCE ADDRESS:
          ADDRESSEE: Human Genome Sciences, Inc.
          STREET: 9410 Key West Avenue
          CITY: Rockville
          STATE: Maryland
          COUNTRY: USA
          ZIP: 20850
     COMPUTER READABLE FORM:
          MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
          COMPUTER: HP Vectra 486/33
          OPERATING SYSTEM: MSDOS version 6.2
          SOFTWARE: ASCII Text
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/08/956,171E
          FILING DATE: 20-Oct-1997
          CLASSIFICATION: <Unknown>
     PRIOR APPLICATION DATA:
          APPLICATION NUMBER: 60/009,861
          FILING DATE: January 5, 1996
          APPLICATION NUMBER: 08/781,986
          FILING DATE: January 3, 1997
     ATTORNEY/AGENT INFORMATION:
          NAME: Mark J. Hyman
          REGISTRATION NUMBER: 46,789
          REFERENCE/DOCKET NUMBER: PB248P1
     TELECOMMUNICATION INFORMATION:
          TELEPHONE: (240) 314-1224
          TELEFAX: (301) 309-8439
     INFORMATION FOR SEQ ID NO: 118:
     SEQUENCE CHARACTERISTICS:
          LENGTH: 4253 base pairs
          TYPE: nucleic acid
          STRANDEDNESS: double
          TOPOLOGY: linear
     SEQUENCE DESCRIPTION: SEQ ID NO: 118:
US-08-956-171E-118

Alignment Scores:
Pred. No.:   Length: 4253
Score: 11.2   Matches: 9
Percent Similarity: 9.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 3.16%   Gaps: 0
                         4

US-09-889-733B-2 (1-285) x US-08-956-171E-118 (1-4253)

Qy   107 GluAlaLeuGluLysLeuGlnLeuAsn 115
            |||||||||||||||||||||||||||
Db   1148 GAAGCATTAGAAAAATTACAATTGAAT 1174

RESULT 15
US-08-286-740-1

; Sequence 1, Application US/08286740
; Patent No. 5561053
; GENERAL INFORMATION:
     APPLICANT: Crowley, Craig W.
```

```
     NUMBER OF SEQUENCES: 4
     CORRESPONDENCE ADDRESS:
          ADDRESSEE: Genentech, Inc.
          STREET: 460 Point San Bruno Blvd
          CITY: South San Francisco
          STATE: California
          COUNTRY: USA
          ZIP: 94080
     COMPUTER READABLE FORM:
          MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
          COMPUTER: IBM PC compatible
          OPERATING SYSTEM: PC-DOS/MS-DOS
          SOFTWARE: patin (Genentech)
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/08/286,740
          FILING DATE: 05-AUG-1994
          CLASSIFICATION: 435
     PRIOR APPLICATION DATA:
          APPLICATION NUMBER:
          FILING DATE:
     ATTORNEY/AGENT INFORMATION:
          NAME: Lee, Wendy M.
          REGISTRATION NUMBER:
          REFERENCE/DOCKET NUMBER: 798
     TELECOMMUNICATION INFORMATION:
          TELEPHONE: 415/225-1994
          TELEFAX: 415/952-9881
          TELEX: 910/371-7168
     INFORMATION FOR SEQ ID NO: 1:
     SEQUENCE CHARACTERISTICS:
          LENGTH: 7360 bases
          TYPE: nucleic acid
          STRANDEDNESS: double
          TOPOLOGY: linear
US-08-286-740-1

Alignment Scores:
Pred. No.:                                Length:        7360
Score:            19.8                     Matches:       9
Percent Similarity:      9.00              Conservative:  0
Best Local Similarity:   100.00%          Mismatches:    0
Query Match:             100.00%          Indels:        0
DB:                      3.16%             Gaps:          0
                         1

US-09-889-733B-2 (1-285) x US-08-286-740-1 (1-7360)

Qy     54   LeuGlySerLysLeuGlyLeuGlnVal 62
            |||||||||||||||||||||||||||
Db     1563 CTCGGTTCTAAGCTTGGGCTGCAGGTC 1589

Search completed: June  8, 2004, 02:59:34
Job time : 111 secs
```

OM protein - nucleic search, using frame_plus_p2n model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: June 8, 2004, 02:18:48 ; Search time 389 Seconds
                                    (without alignments)
                                    3342.336 Million cell updates/sec

Title: US-09-889-733B-2
Perfect score: 285
Sequence: 1 MAGLGHPSAFGRATHAVVRA.........EKVDGLLTCCSVLINKKVDS 285

Scoring table: OLIGO
                Xgapop   60.0 , Xgapext 60.0
                Ygapop   60.0 , Ygapext 60.0
                Fgapop   6.0 , Fgapext   7.0
                Delop   6.0 , Delext   7.0

Searched: 2995936 seqs, 2280998010 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5983782

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_P/US09889733/runat_07062004_083558_8409/runat_07062004_083558_8409/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFIX=olip2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LCOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09889733 @CGN_1_1.723 @runat_07062004_083558_8409
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :   Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result   Score Query   DB ID   Description
No.   Match Length

No.		Score			DB		US-		Sequence Appl
1	259	90.9	1633	17	US-10-450-826-9	Sequence 9, Appli
2	259	90.9	4208	17	US-10-468-125-22	Sequence 22, Appl
3	225	78.9	3778	15	US-10-252-157-328	Sequence 328, App
4	203	71.2	3651	9	US-09-925-300-512	Sequence 512, App
5	106	37.2	3008	15	US-10-205-219-83	Sequence 83, Appli
6	49	17.2	83946	17	US-10-450-826-8	Sequence 8, Appli
7	39	13.7	3189	10	US-09-814-353-20553	Sequence 20553, A
8	20	7.0	60	10	US-09-908-975-9030	Sequence 9030, Ap
9	15	5.3	452	10	US-09-918-995-14871	Sequence 14871, A
10	15	5.3	472	10	US-09-918-995-9117	Sequence 9117, Ap
11	15	5.3	1397	17	US-10-641-643-37	Sequence 37, Appl
12	11	3.9	239	9	US-09-736-457-404	Sequence 404, Appl
13	11	3.9	239	9	US-09-902-941-404	Sequence 404, App
14	11	3.9	239	9	US-09-849-626-404	Sequence 404, App
15	11	3.9	239	10	US-09-476-300-404	Sequence 404, App
16	11	3.9	239	13	US-10-283-017-404	Sequence 404, App
17	11	3.9	239	15	US-10-017-754-404	Sequence 404, App
18	11	3.9	239	15	US-10-113-872-404	Sequence 404, App
19	10	3.5	332	9	US-09-783-590-5309	Sequence 5309, Ap
20	10	3.5	770	9	US-09-910-943-9	Sequence 9, Appli
21	9	3.2	456	9	US-09-815-242-9440	Sequence 9440, Ap
22	9	3.2	640	15	US-10-005-675-18	Sequence 18, Appl
23	9	3.2	645	13	US-10-027-632-215186	Sequence 215186,
24	9	3.2	645	16	US-10-027-632-215186	Sequence 215186,
25	9	3.2	664	15	US-10-005-675-16	Sequence 16, Appl
26	9	3.2	664	15	US-10-005-675-17	Sequence 17, Appl
27	9	3.2	774	15	US-10-288-930-43	Sequence 43, Appl
28	9	3.2	1035	13	US-10-282-122A-11501	Sequence 11501, A
29	9	3.2	4253	8	US-08-781-986A-118	Sequence 118, App
30	9	3.2	4253	13	US-10-329-624-118	Sequence 118, App
31	9	3.2	6491	9	US-09-070-927A-156	Sequence 156, App
32	9	3.2	10011	13	US-10-158-844-76	Sequence 76, Appl
33	9	3.2	69168	13	US-10-087-192-124	Sequence 124, App
34	8	2.8	60	10	US-09-908-975-9031	Sequence 9031, Ap
35	8	2.8	290	13	US-10-424-599-86153	Sequence 86153, A
36	8	2.8	403	10	US-09-918-995-19254	Sequence 19254, A
37	8	2.8	445	10	US-09-918-995-5286	Sequence 5286, Ap
38	8	2.8	453	13	US-10-424-599-21791	Sequence 21791, A
39	8	2.8	471	13	US-10-424-599-120762	Sequence 120762,
40	8	2.8	500	11	US-09-864-408A-8175	Sequence 8175, Ap
41	8	2.8	588	13	US-10-027-632-281162	Sequence 281162,
42	8	2.8	588	16	US-10-027-632-281162	Sequence 281162,
43	8	2.8	628	13	US-10-027-632-12419	Sequence 12419, A
44	8	2.8	628	16	US-10-027-632-12419	Sequence 12419, A
45	8	2.8	632	13	US-10-027-632-104847	Sequence 104847,

ALIGNMENTS

RESULT 1
US-10-450-826-9
; Sequence 9, Application US/1450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
; APPLICANT: Ji, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathon S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Eistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiatic
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9

```
;   ;   LENGTH: 1633
;   ;   TYPE: DNA
;   ;   ORGANISM: Homo sapiens
;   ;   FEATURE:
;   ;   OTHER INFORMATION: Genbank Accession No. AB001915
;   ;   FEATURE:
;   ;   NAME/KEY: unsure
;   ;   LOCATION: (1)..(1633)
;   ;   OTHER INFORMATION: n = a or c or g or t
US-10-450-826-9

Alignment Scores:
Pred. No.:              8.08e-264      Length:             1633
Score:                  259.00         Matches:            259
Percent Similarity:     100.00%        Conservative:       0
Best Local Similarity:  100.00%        Mismatches:         0
Query Match:            90.88%         Indels:             0
DB:                                    Gaps:               0

US-09-889-733B-2 (1-285) x US-10-450-826-9 (1-1633)

Qy      27  GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln  46
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     401  CAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGACGTCGCCCGCGCCGGAACGGCAG  460

Qy      47  HisGlnLeuTyrValGlyValLeuGlySerLysLysGlnLeuValValGlnLeuPro  66
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     461  CACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCCTGCAGGTGGTGCAGCTGCCG  520

Qy      67  AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu  86
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     521  GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAG  580

Qy      87  ThrAlaLeuIleThrArgProGlyAlaProSerArgArgArgValGluValAspMetMetLys  106
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     581  ACGGCCCTCATCACCCGACCCGGGCGCCGGAGCCGGAGGAGGTTGACATGATGAAA  640

Qy     107  GluAlaLeuGluLysGluGlnLeuGlnLeuAsnIleValGluMetLysGluAsnAlaThrLeu  126
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     641  GAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTA  700

Qy     127  AspGlyGlyGlyAlaAspValLeuPheThrGlyArgGluPheGluValGlyLeuSerLysArgThr  146
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     701  GATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCCAAAAGGACA  760

Qy     147  AsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrVal  166
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     761  AATCAACAAGAGGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTG  820

Qy     167  ProValAlaAspGlyValLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle  186
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     821  CCAGTGGCAGATGGGTTGCATTTGAACAGTTTCTGCAGCATGGCTGGCCTAACCTGATC  880

Qy     187  AlaIleGlySerGluSerAlaGlnIleLysAlaLeuLysIleMetGlnGlnMetSerAsp  206
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     881  GCAATTGGGTCTGAGTCAGCTCAGATCAAAGCCCTTAAGATCATGCAACAGATGAGTGAC  940

Qy     207  HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAlaAsnCysIleTyrLeuAsn  226
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     941  CACCGCTACGACAAACTCACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAAT  1000

Qy     227  IleProAsnLysGlyHisHisArgThrProGluGluGlnTyrProGluSerAla  246
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1001  ATCCCCAACAAAGGGCACGTCTTGCTCACCGGAACCCCGGAAGAGTATCCAGAAAGTGCA  1060

Qy     247  LysValTyrGluValLeuLysAspHisMetLeuIleProValSerMetSerGluGluGlu  266
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1061  AAGGTTTATGAAGTACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAA  1120

Qy     267  LysValAspGlyValLeuGluThrCysCysSerValLeuIleAsnLysLysValAspSer  285
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1121  AAGGTGGATGGCGTGCTCGAGACCTGCTGCTCAGTTTTAATTAACAAGAAAGTAGACTCC  1177
```

```
RESULT 2
US-10-468-125-22
;   ;   Sequence 22, Application US/10468125
;   ;   Publication No. US20040082061A1
;   ;   GENERAL INFORMATION:
;   ;   APPLICANT: ASTROMOFF, Anna
;   ;   APPLICANT: AU-YOUNG, Janice
;   ;   APPLICANT: BAUGHN, Mariah R.
;   ;   APPLICANT: DING, Li
;   ;   APPLICANT: DUGGAN, Brendan M.
;   ;   APPLICANT: FORSYTHE, Ian J.
;   ;   APPLICANT: GIETZEN, Kimberly J.
;   ;   APPLICANT: GRIFFIN, Jennifer A.
;   ;   APPLICANT: LEE, Ernestine A.
;   ;   APPLICANT: LU, Yan
;   ;   APPLICANT: RICHARDSON, Thomas W.
;   ;   APPLICANT: RING, Huijun Z.
;   ;   APPLICANT: SANJANWALA, Madhusudan
;   ;   APPLICANT: SWARNAKAR, Anita
;   ;   APPLICANT: CHAWLA, Narinder K.
;   ;   APPLICANT: WARREN, Bridget A.
;   ;   APPLICANT: XU, Yuming
;   ;   APPLICANT: YUE, Henry
;   ;   APPLICANT: ZEBARJADIAN, Yeganeh
;   ;   TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
;   ;   FILE REFERENCE: PI-0363 USN
;   ;   CURRENT APPLICATION NUMBER: US/10/468,125
;   ;   CURRENT FILING DATE:  2003-08-15
;   ;   PRIOR APPLICATION NUMBER: PCT/US02/04918
;   ;   PRIOR FILING DATE: 2002-02-14
;   ;   PRIOR APPLICATION NUMBER: US 60/269,643
;   ;   PRIOR FILING DATE: 2001-02-16
;   ;   PRIOR APPLICATION NUMBER: US 60/271,332
;   ;   PRIOR FILING DATE: 2001-02-23
;   ;   PRIOR APPLICATION NUMBER: US 60/276,767
;   ;   PRIOR FILING DATE: 2001-03-16
;   ;   PRIOR APPLICATION NUMBER: US 60/282,077
;   ;   PRIOR FILING DATE: 2001-04-06
;   ;   PRIOR APPLICATION NUMBER: US 60/285,447
;   ;   PRIOR FILING DATE: 2001-04-19
;   ;   PRIOR APPLICATION NUMBER: US 60/287,060
;   ;   PRIOR FILING DATE: 2001-04-27
;   ;   PRIOR APPLICATION NUMBER: US 60/288,543
;   ;   PRIOR FILING DATE: 2001-05-03
;   ;   NUMBER OF SEQ ID NOS: 24
;   ;   SEQ ID NO 22
;   ;   SOFTWARE: PERL Program
;   ;   LENGTH: 4208
;   ;   TYPE: DNA
;   ;   ORGANISM: Homo sapiens
;   ;   FEATURE:
;   ;   NAME/KEY: misc_feature
;   ;   OTHER INFORMATION: Incyte ID No: 7671089CB1
US-10-468-125-22

Alignment Scores:
Pred. No.:              1.89e-263      Length:             4208
Score:                  259.00         Matches:            259
Percent Similarity:     100.00%        Conservative:       0
Best Local Similarity:  100.00%        Mismatches:         0
Query Match:            90.88%         Indels:             0
DB:                                    Gaps:               0

US-09-889-733B-2 (1-285) x US-10-468-125-22 (1-4208)

Qy      27  GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln  46
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     446  CAGCACGCGCTGAGAAGCGCCAAGGGCGAGGAGGTGACGTCGCCCGCGCCGGAACGGCAG  505

Qy      47  HisGlnLeuTyrValGlyValLeuGlySerLysLysGlnLeuValValGlnLeuPro  66
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     506  CACCAGCTCTACGTGGGCGTGCTGGGCAGCAAGCTGGGCCTGCAGGTGGTGCAGCTGCCG  565
```

```
Pred. No.:              3778
Score:                  225.00
Percent Similarity:     99.58%
Best Local Similarity:  99.58%
Query Match:            78.95%
DB:
Length:        3778
Matches:        238
Conservative:     0
Mismatches:       0
Indels:           1
Gaps:             0
1.57e-227
```

US-09-889-733B-2 (1-285) x US-10-252-157-328 (1-3778)

```
Qy   48 GlnLeuTyrValGlyValGlyLeuGlyLeuGlnValValGluGlnProAla       67
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 CAGCTCTACGTGGGCGGTGCTGGGCAGCAAGCGTGGGTGCCAGGTGGTGGAGCAGCCGCC  61

Qy   68 AspGluSerLeuProAspCysValPheValGluAlaValAlaValCysGluGluThr      87
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 GACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGTGTGCGAGGAGACG  121

Qy   88 AlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLysGlu     107
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 GCCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAAGAA  181

Qy  108 AlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAsp     127
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182 GCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAGATGAAAATGCAACTTTAGAT  241

Qy  128 GlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThrAsn     147
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242 GGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGGGCTTTCAAAAGGACAAAT  301

Qy  148 -GlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysAspThryrAlaValSerThrValPr    167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302 CCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTGCC  361

Qy  167 oValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAl    187
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  362 AGTGGCAGATGGGTTGCATTGAAGAGTTTCTGCAGCATGCCTAACCTGATCGC  421

Qy  187 aIleGlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHi    207
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  422 AATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCA  481

Qy  207 sArgTyrAspLysProAspAspIleAlaLeuAlaAsnCysIleTyrLeuAsnIl    227
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  482 CCGCTACGACAAACTCACTGTGCCTGATGACATAGCACAAACTGTATATATCTAAATAT  541

Qy  227 eProAsnLysGlyHisValLeuLeuHisArgThrProGluTyrProGluSerAlaLy    247
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  542 CCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGAAGAGTATCCAGAAAGTGCAAA  601

Qy  247 sValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLy    267
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  602 GGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAA  661

Qy  267 sValAspLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer    285
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  662 GGTGGATGGGCTGCTCCACCTGCTGCTCAGTTTTAATTAACAAGAAAGTAGACTCC  716
```

```
RESULT 4
US-09-925-300-512
; Sequence 512, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT: Steve Ruben
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA101
;  CURRENT APPLICATION NUMBER: US/09/925,300
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05988
; ,PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1890
```

```
Qy   67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValAlaValCysGluGlu  86
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  566 GCCGACGAGAGCCTTCCGGACTGCGTCTTCGTGGAGGACGTGGCCGTGTGTGCGAGGAG  625

Qy   87 ThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLys     106
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  626 ACGGCCCTCATCACCCGACCCGGGGCGCCGAGCCGGAGGAAGGAGGTTGACATGATGAAA  685

Qy  107 GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu     126
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  686 GAAGCATTAGAAAAACTTCAGCTCAATATAGTAGAGATGAAGATGAAAATGCAACTTTA  745

Qy  127 AspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThr     146
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  746 GATGGCGGAGATGTTTTATTCACAGGCAGAGAATTTTTGTGGGCCTTTCAAAAGGACA  805

Qy  147 AsnGlnArgGlyAlaGluIleLeuLeuAlaAspThrPheLysAspThrTyrAlaValSerThrVal  166
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  806 AATCAACGAGGTGCTGAAATCTTGGCTGATACTTTAAGGACTATGCAGTCTCCACAGTG  865

Qy  167 ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle     186
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  866 CCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGCCTGGCCTAACCTGATC  925

Qy  187 AlaIleGlySerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp     206
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  926 GCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGAC  985

Qy  207 HisArgTyrAspLysProAspAspIleAlaLeuAlaAsnCysIleTyrLeuAsn     226
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  986 CACCGCTACGACAAACTCACTGTGCCTGATGACATAGCACAAACTGTATATATCTAAAT 1045

Qy  227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluTyrProGluSerAla     246
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1046 ATCCCCAACAAAGGGCACGTCTTGCTGCACCGAACCCCGAAGAGTATCCAGAAAGTGCA 1105

Qy  247 LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu     266
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1106 AAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAA 1165

Qy  267 LysValAspLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer     285
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1166 AAGGTGGATGGGCTGCTCCACCTGCTGCTCAGTTTTAATTAACAAGAAAGTAGACTCC 1222
```

```
RSSULT 3
US-10-252-157-328
; Sequence 328, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
;  APPLICANT: Faris, Mary
;  APPLICANT: Pearson, Cecelia I.
;  TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
;  FILE REFERENCE: PA-0027-1 US
;  CURRENT APPLICATION NUMBER: US/10/252,157
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: 60/295,048
;  PRIOR FILING DATE: 2001-05-31
;  NUMBER OF SEQ ID NOS: 501
;  SOFTWARE: PERL Program
; SEQ ID NO 328
;  LENGTH: 3778
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc feature
;  OTHER INFORMATION: Incyte ID No. US2003019064OA1 198087.1
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 1083-1106
;  OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-328

Alignment Scores:
```

```
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 512
;   LENGTH: 3651
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (1283)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (3641)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (3650)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-512

Alignment Scores:
Pred. No.:                  2.85e-204        Length:         3651
Score:                      203.00           Matches:        203
Percent Similarity:         100.00%          Conservative:   0
Best Local Similarity:      100.00%          Mismatches:     0
Query Match:                71.23%           Indels:         0
DB:                                          Gaps:           0

US-09-889-733B-2 (1-285) x US-09-925-300-512 (1-3651)

Qy   73  AspCysValPheValGluArgValAlaValValCysGluThrAlaLeuIleThrArg   92
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4  GACTGCGCTTCGTGCGTGGAGAGGACGTGGCCGTGGTGTGCGAGGAGACGGCCCTCATCACCCGA   63

Qy   93  ProGlyAlaProSerArgArgLysGluValAspMetMetLysGluAlaLeuGluLysLeu  112
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64  CCCGGGGCCGCCGAGCCCGGAGAAGAGAGGTTGACATGATGAAAGAAGCATTAGAAAAACTT  123

Qy  113  GlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeu  132
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124  CAGCTCAATATAGTAGAGATGAAAGATGAAAATGCAACTTTAGATGGCGGAGATGTTTTA  183

Qy  133  PheThrGlyArgGluPheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGlu  152
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  184  TTCACAGGCAGAGAATTTTGTGGGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAA  243

Qy  153  IleLeuHisAspThrPheLysAspTyrAlaValSerThrValProValAlaAspGlyLeu  172
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  244  ATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTG  303

Qy  173  HisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGlu  192
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  304  CATTTGAAGAGTTTCTGCAGCATGGCTGGGCCTAACCTGATCGCAATTGGTTCTAGTGAA  363

Qy  193  SerAlaGlnLysLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLysSerLeu  212
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  364  TCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGACCACCGCTACGACAAACTC  423

Qy  213  ThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHis  232
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  424  ACTGTGCCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCCAACAAAGGGCAC  483

Qy  233  ValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLysLeu  252
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  484  GTCTTGCTGCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAAGGTTTATGAGAAACTG  543

Qy  253  LysAspHisMetLeuIleProValSerMetSerGluLeuGluLysMetGluValAspGlyLeuLeu  272
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  544  AAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGATGGAAGTGGATGGGCTC  603

Qy  273  ThrCysCys  275
         |||||||
Db  604  ACCTGCTGT  612

RESULT 5
US-10-205-219-83
```

```
;   Sequence 83, Application US/10205219
;   Publication No.US20030138803A1
;   GENERAL INFORMATION:
;   APPLICANT: Warner-Lambert Company
;   APPLICANT: Lee, Kevin.
;   APPLICANT: Dixon, Alistair
;   APPLICANT: Brooksbank, Robert
;   APPLICANT: Pinnock, Robert
;   TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
;   FILE REFERENCE: WL-A-018200
;   CURRENT APPLICATION NUMBER: US/10/205,219
;   CURRENT FILING DATE: 2002-07-24
;   PRIOR APPLICATION NUMBER: GB 0118354.0
;   PRIOR FILING DATE: 2001-07-27
;   NUMBER OF SEQ ID NOS: 197
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 83
;   LENGTH: 3008
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: N-G, N-G dimethylarginine dimethylaminohydrolase
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 2411_2412
;   OTHER INFORMATION: n is a or g or c or t
US-10-205-219-83

Alignment Scores:
Pred. No.:                  9.57e-102        Length:         3008
Score:                      106.00           Matches:        152
Percent Similarity:         98.70%           Conservative:   0
Best Local Similarity:      98.70%           Mismatches:     1
Query Match:                37.19%           Indels:         2
DB:                                          Gaps:           0

US-09-889-733B-2 (1-285) x US-10-205-219-83 (1-3008)

Qy   65  LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaAlaValValCys   84
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  624  CTGCCGGCCGACGAGAGCCTGCCTGACTGCGTGTTCGTGGAGGACGTGGCCGTGGTGTGC  683

Qy   85  GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet  104
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  684  GAGGAGACCGCCCTCATCACCCGCCCGGGCGCCCTAGCCGCAGGAAGGAGGTTGACATG  743

Qy  105  MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla  124
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  744  ATGAAAGAGGCTTTGGAAAAACTTCAGCTCAACATAGTAGAGATGAAAGATGAAAATGCA  803

Qy  125  ThrLeuAspGlyGlyAspValPheThrGlyArgGluPheValGlyLeuSerLys  144
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  804  ACCTTAGATGGTGGGGACGTCCTATTCACAGGCAGAGAGTTTTTGTGGGCCTTTCCAAA  863

Qy  145  ArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSer  164
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  864  AGGACAAATCAACAGGGTGCTGAGATCTTGGCTGATACTTTCAAGGACTACGCAGTTTCC  923

Qy  165  ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs  184
         ||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  924  ACAGTCCCCGTGGCCGATTC-TTTGCATTTAAAGAGTTTCTGCAGCATGCTGGCCCAA  982

Qy  184  nLeuIleAlaIleGlySerSerGluSerAlaGlnLysLysAlaLeuLysIleMetGlnGlnMe  204
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  983  CCTGATCGCAATAGGGTCCAGTGAATCTGCCCAGAAGCCCTCAAGATCATGCAACAGAT  1042

Qy  204  tSerAspHisArgTyrAspLysThrValProAspAsp  217
         |||||||||||||||||||||||||||||||||||
Db 1043  GAGTGACCACCGTTATGACAAGCTCACTGTACCGACGAC  1082

RESULT 6
US-10-450-826-8/c
;   Sequence 8, Application US/10450826
```

Publication No. US20040101818A1
GENERAL INFORMATION:
; APPLICANT: Ji, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathon S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Einstein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiati
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 83946
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. AL078459
US-10-450-826-8

Alignment Scores:
Pred. No.:   3.68e-40   Length: 83946
Score: 49.00   Matches: 49
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 17.19%   Indels: 0
DB: 17   Gaps: 0

US-09-889-733B-2 (1-285) x US-10-450-826-8 (1-83946)

Qy   200 IleMetGlnGlmMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAla 219
Db 18610 ATCATGCAACAGATGAGTGACCACCGCTACGACAAACTCACTGTGCCTGATGACATAGCA 18551

Qy   220 AlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro 239
Db 18550 GCAAACTGTATATATCTAAATATCCCCAACAAAGGGCACGTTCTTGCTGCACCGAACCCCG 18491

Qy   240 GluGluTyrProGluSerAlaLysVal 248
Db 18490 GAAGAGTAATCCAGAAAGTGCAAAGGTA 18464

RESULT 7
US-09-814-353-20553/c
; Sequence 20553, Application US/09814353
; Publication No. US20030165831A1
GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25

; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20553
;   LENGTH: 3189
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1, 2_3189
;   OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20553

Alignment Scores:
Pred. No.:   7.46e-31   Length: 3189
Score: 39.00   Matches: 39
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 13.68%   Indels: 0
DB: 10   Gaps: 0

US-09-889-733B-2 (1-285) x US-09-814-353-20553 (1-3189)

Qy   247 LysValTyrGluLysGluValProValSerMetSerGluLeuGlu 266
Db   3186 AAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAA 3127

Qy   267 LysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer 285
Db   3126 AAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTAATTAACAAGAAAGTAGACTCC 3070

RESULT 8
US-09-908-975-9030
; Sequence 9030, Application US/09908975
; Publication No. US20030165843A1
GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE \
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9030
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-9030

Alignment Scores:
Pred. No.:   2.67e-12   Length: 60
Score: 20.00   Matches: 20
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 7.02%   Indels: 0
DB: 10   Gaps: 0

US-09-889-733B-2 (1-285) x US-09-908-975-9030 (1-60)

Qy   251 LysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGly 270
Db   1 AAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGG 60

RESULT 9
US-09-918-995-14871
; Sequence 14871, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;   APPLICANT: Hyseq, Inc.
;   TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;   TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;   FILE REFERENCE: 20411-756
;   CURRENT APPLICATION NUMBER: US/09/918,995
;   CURRENT FILING DATE: 2001-07-30
;   PRIOR APPLICATION NUMBER: US/09/235,076
;   PRIOR FILING DATE: 1999-01-20
;   NUMBER OF SEQ ID NOS: 38054
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14871
;   LENGTH: 452
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(452)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-14871

Alignment Scores:
Pred. No.:   3.16e-06   Length: 452
Score: 15.00   Matches: 15
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 5.26%   Indels: 0
DB: 10   Gaps: 0

US-09-889-733B-2 (1-285) x US-09-918-995-14871 (1-452)

Qy   130 AspValLeuPheThrGlyArgGluPheValGlyLeuSerLys 144
         ||||||||||||||||||||||||||||||||||||||||||
Db   324 GACGTTCTCTTCACCGGCCGCGGAGTTTTCGTAGGCCTCCAAA 368

US-09-889-733B-2 (1-285) x US-09-918-995-14871

RESULT 10
US-09-918-995-9117
; Sequence 9117, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;   APPLICANT: Hyseq, Inc.
;   TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;   TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;   FILE REFERENCE: 20411-756
;   CURRENT APPLICATION NUMBER: US/09/918,995
;   CURRENT FILING DATE: 2001-07-30
;   PRIOR APPLICATION NUMBER: US/09/235,076
;   PRIOR FILING DATE: 1999-01-20
;   NUMBER OF SEQ ID NOS: 38054
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9117
;   LENGTH: 472
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(472)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-9117

Alignment Scores:
Pred. No.:   3.29e-06   Length: 472
Score: 15.00   Matches: 15
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 5.26%   Indels: 0
DB: 10   Gaps: 0

US-09-889-733B-2 (1-285) x US-09-918-995-9117 (1-472)

Qy   130 AspValLeuPheThrGlyArgGluPheValGlyLeuSerLys 144
         ||||||||||||||||||||||||||||||||||||||||||
Db   324 GACGTTCTCTTCACCGGCCGCGGAGTTTTCGTAGGCCTCCAAA 368

Qy   130 AspValLeuPheThrGlyArgGluPheValGlyLeuSerLys 144
         ||||||||||||||||||||||||||||||||||||||||||
Db   357 GACGTTCTCTTCACCGGCCGCGGAGTTTTCGTAGGCCTCCAAA 401

RESULT 11
US-10-641-643-37
; Sequence 37, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
;   APPLICANT: Cocks, Benjamin G.
;   Susan G. Stuart
;   Jeffrey J. Seilhamer
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
;   GENE EXPRESSION
;   NUMBER OF SEQUENCES: 1508
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/641,643
;   FILING DATE: 14-Aug-2003
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: <Unknown>
;   FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 37:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1397 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: HMC1NOT01
;   CLONE: 003437
;   SEQUENCE DESCRIPTION: SEQ ID NO: 37 :
US-10-641-643-37

Alignment Scores:
Pred. No.:   8.68e-06   Length: 1397
Score: 15.00   Matches: 15
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 5.26%   Indels: 0
DB: 17   Gaps: 0

US-09-889-733B-2 (1-285) x US-10-641-643-37 (1-1397)

Qy   130 AspValLeuPheThrGlyArgGluPheValGlyLeuSerLys 144
         ||||||||||||||||||||||||||||||||||||||||||
Db   716 GACGTTCTCTTCACCGGCCGCGGAGTTTTCGTAGGCCTCCAAA 760

RESULT 12
US-09-736-457-404/c
; Sequence 404, Application US/09736457
; Patent No. US2002016837A1
; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-404

Alignment Scores:
Pred. No.:   0.0304   Length: 239
Score: 11.00   Matches: 11
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 3.86%   Indels: 0
DB: 9   Gaps: 0

US-09-889-733B-2 (1-285) x US-09-736-457-404 (1-239)

Qy   130 AspValLeuPheThrGlyArgGluPhePheVal 140
          |||||||||||||||||||||||||||||
Db   35 GACGTTCTCTTCACCGGCCGCGGGAGTTTTCGTA 3

RESULT 13
US-09-902-941-404/c
; Sequence 404, Application US/09902941
; Patent No. US2002017295A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-404

Alignment Scores:
Pred. No.:   0.0304   Length: 239
Score: 11.00   Matches: 11
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 3.86%   Indels: 0
DB: 9   Gaps: 0

US-09-889-733B-2 (1-285) x US-09-902-941-404 (1-239)

Qy   130 AspValLeuPheThrGlyArgGluPhePheVal 140
          |||||||||||||||||||||||||||||
Db   35 GACGTTCTCTTCACCGGCCGCGGGAGTTTTCGTA 3

RESULT 14
US-09-849-626-404/c
; Sequence 404, Application US/0984962
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-626-404

Alignment Scores:
Pred. No.:   0.0304   Length: 239
Score: 11.00   Matches: 11
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 3.86%   Indels: 0
DB: 9   Gaps: 0

US-09-889-733B-2 (1-285) x US-09-849-626-404 (1-239)

Qy   130 AspValLeuPheThrGlyArgGluPhePheVal 140
          |||||||||||||||||||||||||||||
Db   35 GACGTTCTCTTCACCGGCCGCGGGAGTTTTCGTA 3

RESULT 15
US-09-476-300-404/c
; Sequence 404, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-476-300-404

Alignment Scores:
Pred. No.:   0.0304   Length: 239
Score: 11.00   Matches: 11
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 3.86%   Indels: 0
DB: 10   Gaps: 0

US-09-889-733B-2 (1-285) x US-09-476-300-404 (1-239)

Qy   130 AspValPheThrGlyArgGluPhePheVal 140
          ||||||||||||||||||||||||||||||
Db   35 GACGTTCTCTTCACCGGCCGGGAGTTTTCGTA 3

Search completed: June 8, 2004, 03:06:40
Job time : 423 secs